(12) **United States Patent**
Eldering et al.

(10) Patent No.: **US 6,457,010 B1**
(45) Date of Patent: ***Sep. 24, 2002**

Profile
creation

(54) **CLIENT-SERVER BASED SUBSCRIBER CHARACTERIZATION SYSTEM**

(75) Inventors: **Charles A. Eldering**, Doylestown, PA (US); **M. Lamine Sylla**, New Britain, PA (US)

(73) Assignee: **Expanse Networks, Inc.**, Doylestown, PA (US)

(*) Notice: Subject to any disclaimer, the term of this patent is extended or adjusted under 35 U.S.C. 154(b) by 0 days.

This patent is subject to a terminal disclaimer.

(21) Appl. No.: **09/205,653**

(22) Filed: **Dec. 3, 1998**

(51) Int. Cl.^{7} G06F 17/00; G06F 17/30

(52) U.S. Cl. 707/10; 707/102; 705/14; 725/14

(58) Field of Search 707/102, 10, 1, 707/2; 345/327; 348/10; 725/10, 14; 705/14

(56) **References Cited**

U.S. PATENT DOCUMENTS

4,573,072 A	*	2/1986	Freeman	348/10
4,602,279 A		7/1986	Freeman	358/86
4,745,549 A		5/1988	Hashimoto	364/402
5,099,319 A		3/1992	Esch et al.	358/86
5,155,591 A		10/1992	Wachob	358/86
5,231,494 A		7/1993	Wachob	358/146
5,251,324 A	*	10/1993	McMullan, Jr.	455/2
5,319,455 A		6/1994	Hoarty et al.	348/7
5,351,075 A		9/1994	Herz et al.	348/1

(List continued on next page.)

FOREIGN PATENT DOCUMENTS

CA	2286243	4/2001
CA	2323166	4/2001
GB	2348346	9/2000
GB	2348530	10/2000
WO	9707565	2/1997
WO	9712486	4/1997
WO	9717774	5/1997
WO	9741673	11/1997

(List continued on next page.)

OTHER PUBLICATIONS

IEEE Publication ""Goodies" in exchange for consumer information on the internet: economics and issues", by Chang, A.M.; Kannan, P.K. and Whinston, A.B., Jan. 1998, 12 pages.
"AdForce Introduces AdForce EveryWhere: Advertising, Marketing & Promotions wherever a Digital Signal Can be Sent", by Business Wire, Mar. 27, 2000 (3pp.).

(List continued on next page.)

Primary Examiner—Safet Metjahic
Assistant Examiner—Uyen Le
(74) *Attorney, Agent, or Firm*—Douglas J. Ryder

(57) **ABSTRACT**

A subscriber characterization system is presented in which the subscriber's requests are transmitted to a server which fulfills those requests and performs monitoring of the subscriber requests for subsequent characterization of the subscriber. Monitoring includes maintaining records of the time duration programming is watched, the volume at which the programming is listened to, and any available information regarding the type of programming, including category and sub-category of the programming. The characterization system works across a network to extract textual information related to the programming from closed captioning data, electronic program guides, or other text sources associated with the programming. The extracted information is used to form program characteristics vectors. The programming characteristics vectors can be used in combination with the subscriber selection data to form a subscriber profile. Heuristic rules indicating the relationships between programming choices and demographics can be applied to generate additional probabilistic information regarding demographics and programming and product interests. The probabilistic information can be accessed at the server by other entities on the network.

48 Claims, 21 Drawing Sheets
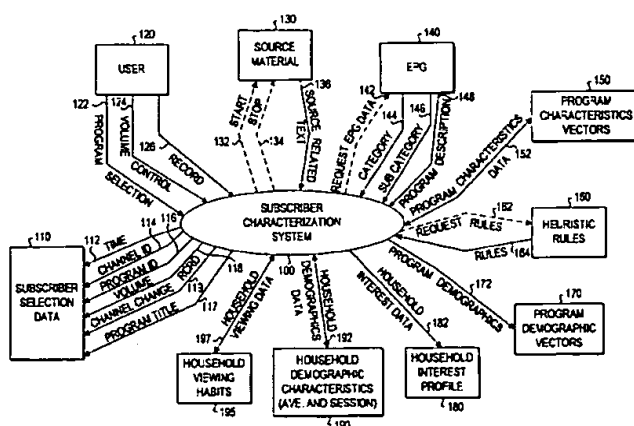
U.S. PATENT DOCUMENTS

5,410,344	A		4/1995	Graves et al. 348/1
5,446,919	A		8/1995	Wilkins 455/6.2
5,559,549	A		9/1996	Hendricks et al. 348/6
5,565,909	A		10/1996	Thibadeau et al. 348/9
5,585,865	A		12/1996	Amano et al. 348/731
5,604,542	A		2/1997	Dedrick 348/552
5,632,007	A		5/1997	Freeman 395/75
5,636,346	A		6/1997	Saxe 395/201
5,661,516	A		8/1997	Carles 348/8
5,710,884	A		1/1998	Dedrick 395/200.47
5,724,521	A		3/1998	Dedrick 395/226
5,740,549	A		4/1998	Reilly et al. 705/14
5,749,081	A		5/1998	Whiteis 707/102
5,754,938	A	*	5/1998	Herz et al. 455/4.2
5,758,257	A		5/1998	Herz et al. 455/2
5,758,259	A		5/1998	Lawler 455/5.1
5,761,601	A		6/1998	Nemirofsky et al. 455/3.1
5,761,662	A		6/1998	Dasan 707/10
5,774,170	A		6/1998	Hite et al. 348/9
5,790,935	A	*	8/1998	Payton 455/5.1
5,796,952	A		8/1998	Davis et al. 395/200.54
5,848,396	A		12/1998	Gerace 705/10
5,912,696	A		6/1999	Buehl 348/5.5
5,915,243	A		6/1999	Smolen 705/14
5,918,014	A		6/1999	Robinson 395/200.49
5,948,061	A		9/1999	Merriman et al. 709/219
5,956,637	A	*	9/1999	Ericsson et al. 455/414
5,974,396	A		10/1999	Anderson et al. 705/10
5,977,964	A	*	11/1999	Williams et al. 345/327
5,991,735	A		11/1999	Gerace 705/10
6,005,597	A		12/1999	Barrett et al. 348/1
6,009,410	A		12/1999	LeMole et al. 705/14
6,012,051	A		1/2000	Sammon, Jr. et al. 706/52
6,020,883	A		2/2000	Herz et al. 345/327
6,055,510	A		4/2000	Henrick et al. 705/14
6,088,722	A		7/2000	Herz et al. 709/217
6,108,637	A		8/2000	Blumenau 705/7
6,119,098	A		9/2000	Guyot et al. 705/14
6,133,912	A	*	10/2000	Montero 345/327
6,134,532	A		10/2000	Lazarus et al. 705/14
6,141,010	A		10/2000	Hoyle 345/356
6,160,570	A	*	12/2000	Sitnik 348/1
6,160,989	A		12/2000	Hendricks et al. 455/4.2
6,177,931	B1		1/2001	Alexander et al. 345/327
6,286,005	B1	*	9/2001	Cannon 707/100
6,327,574	B1	*	12/2001	Kramer et al. 705/14

FOREIGN PATENT DOCUMENTS

WO	9821877	5/1998
WO	9828906	7/1998
WO	9834189	8/1998
WO	9901984	1/1999
WO	9904561	1/1999
WO	9904166	9/1999
WO	9944159	9/1999
WO	9965237	12/1999
WO	9966719	12/1999
WO	0008802	2/2000
WO	0013434	3/2000
WO	0014951	3/2000
WO	0022818	4/2000
WO	0033224	6/2000
WO	0049801	8/2000
WO	0054504	9/2000
WO	0055748	9/2000
WO	0124087	4/2001

OTHER PUBLICATIONS

"The AdForce professional management team, composed of seasoned high technology and Internet executives, is unique in the field of Internet advertising management.", by AdForce Co., Mar. 27, 2000. (5 pp.).

"AdForce Strategic Partners", by AdForce Co., Mar. 27, 2000. (3 pp.).

"Data Structures: From Arrays to Priority Queues", by Wayne Amsbury, Wadsworth Publishing Company, 1985, pp. 228, 332 and 333. (3pp.).

IEEE Publication "A Framework for Targeting Banner Advertising on the Internet", by Gallagher, K. and Parsons, J., Jan. 1997, 12 pages.
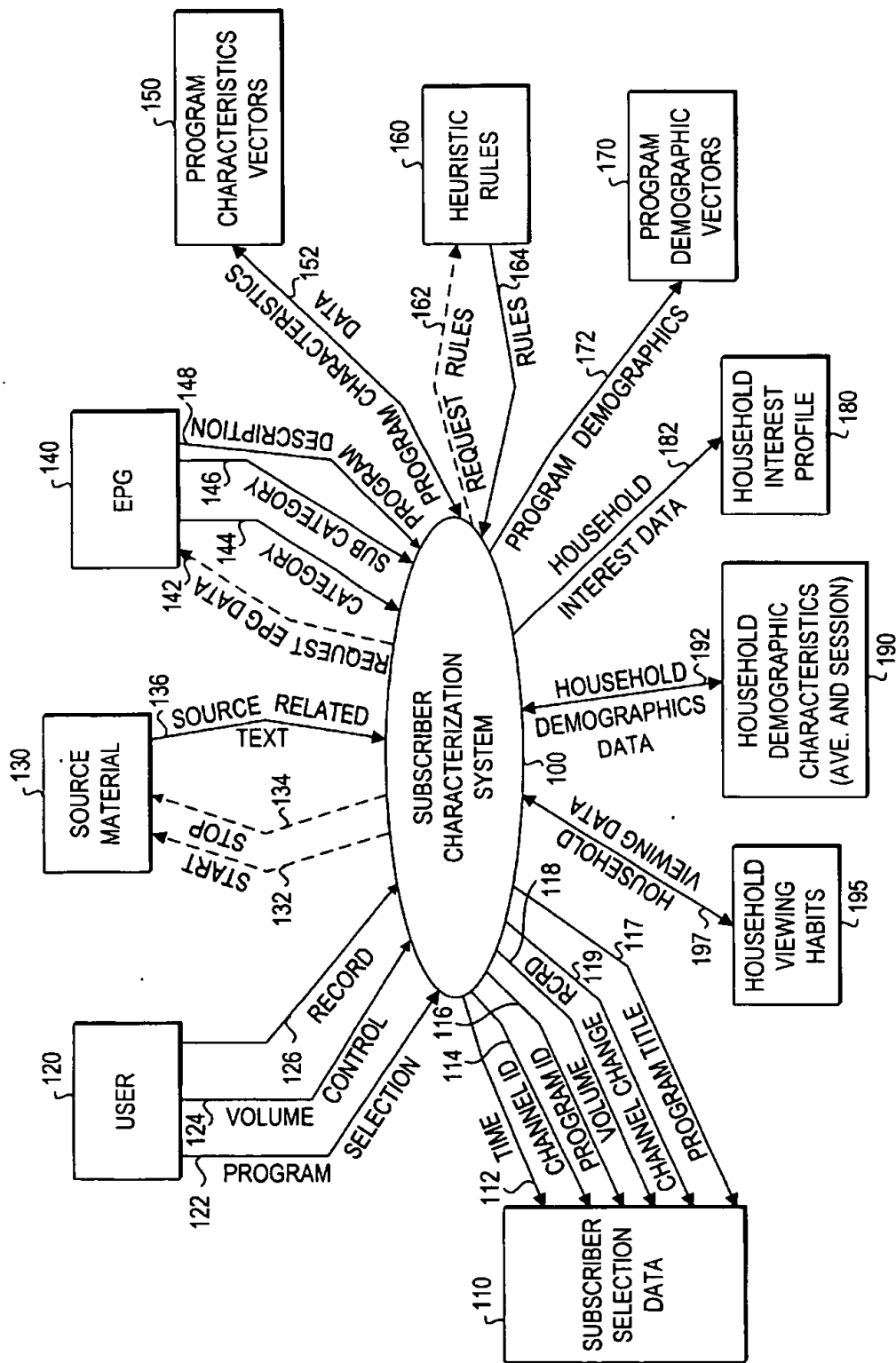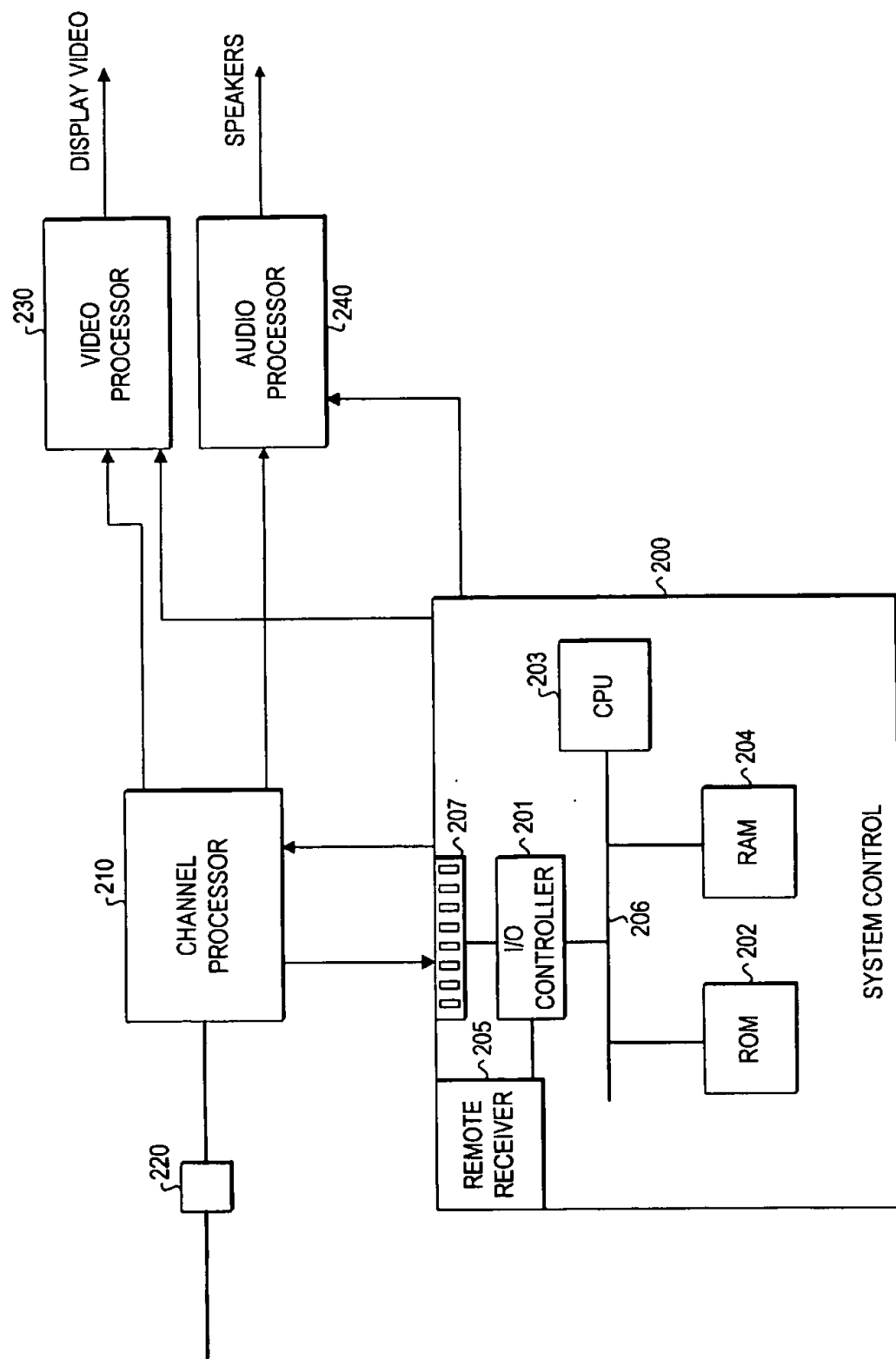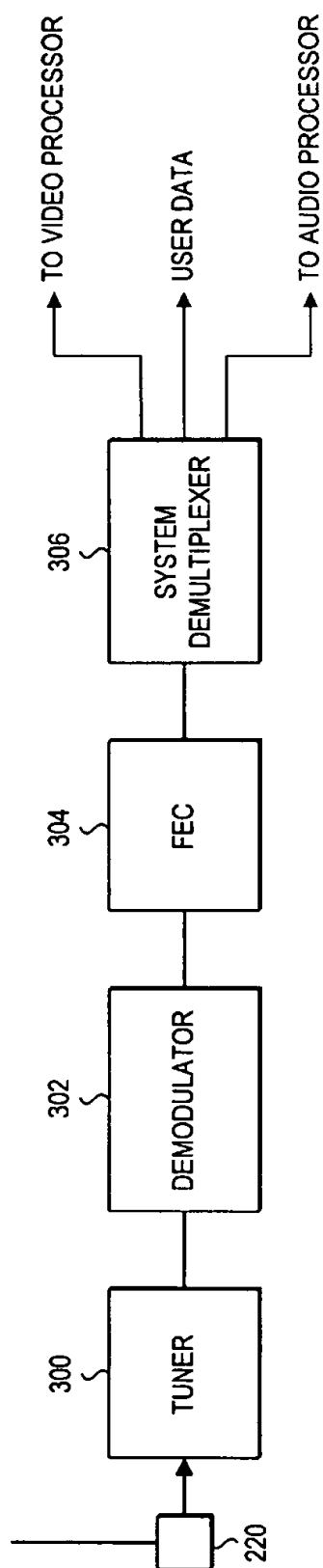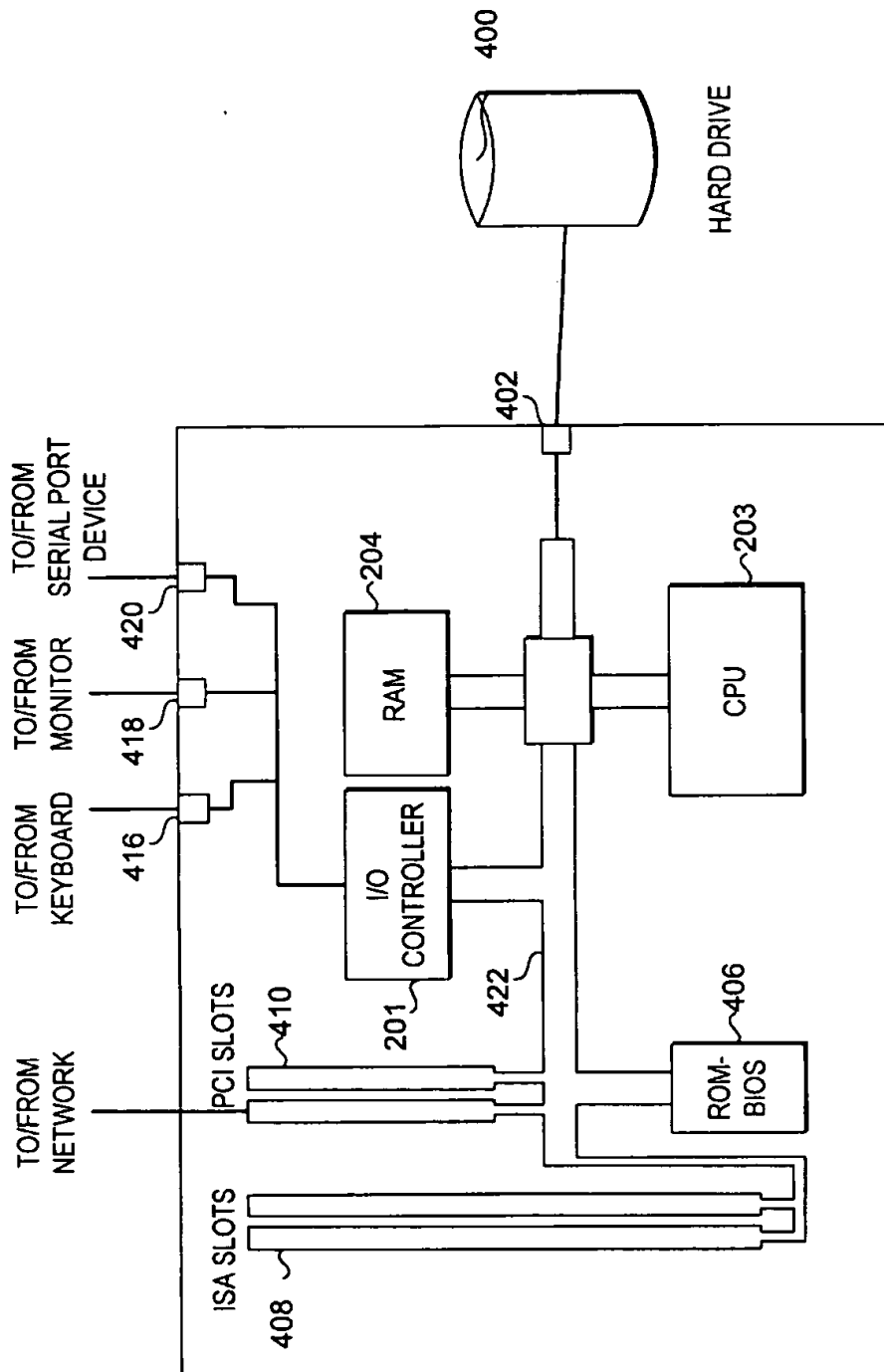
* cited by examiner

FIG. 1

FIG. 2

FIG. 3

FIG. 4
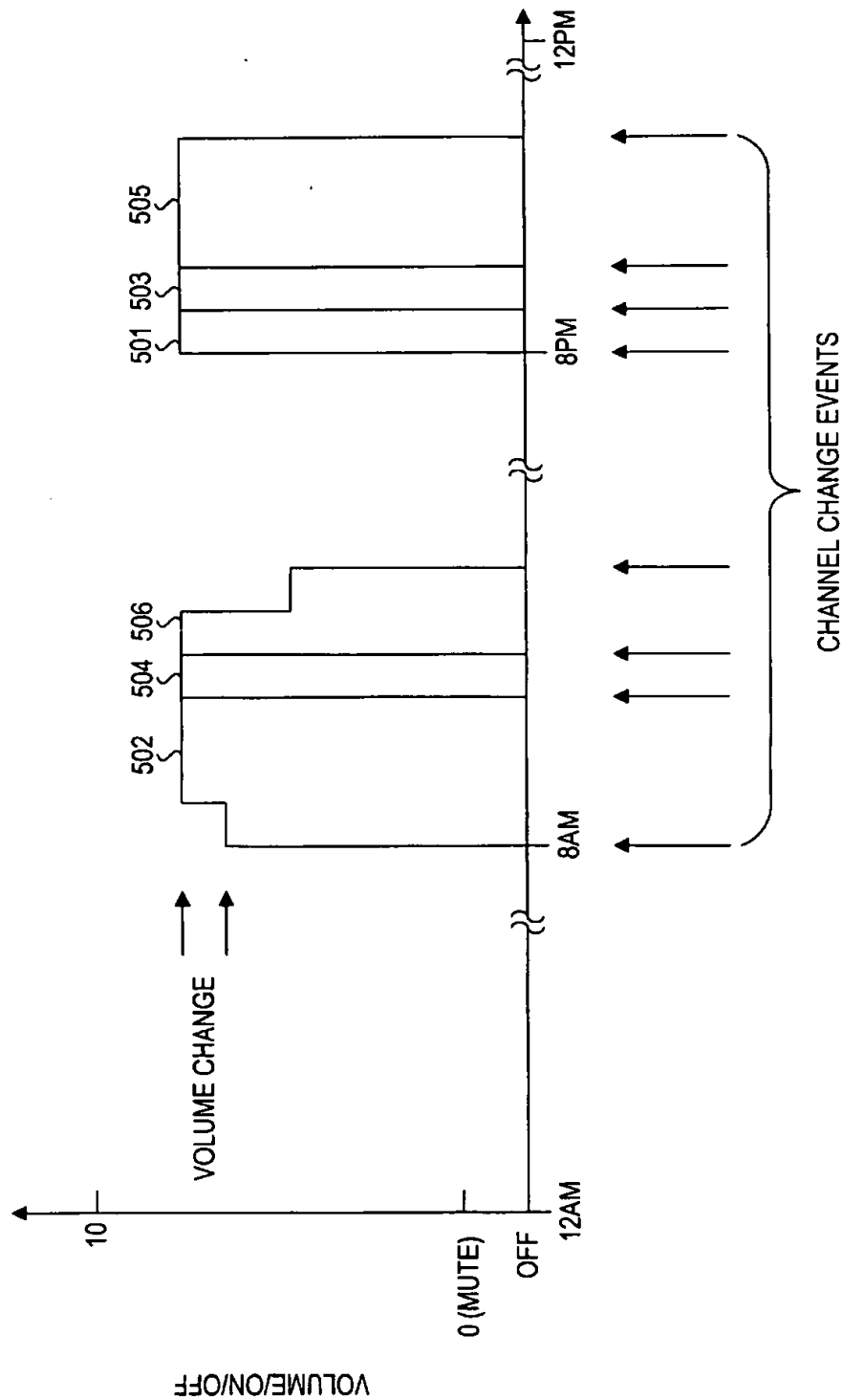
FIG. 5

TIME	CHANNEL ID	PROGRAM TITLE	VOLUME
08:01:25AM	06	"MORNING TV"	5/10
08:01:45AM	13	"GOOD MORNING AMERICA"	5/10
08:03:25AM	13	"GOOD MORNING AMERICA"	6/10
...			
06:11:25PM	09	"SEINFELD"	5/10
06:15:23PM	09	"ADVERTISING"	5/10
06:17:25PM	09	"SEINFELD"	5/10
06:28:10PM	09	"ADVERTISING"	5/10
06:30:07PM	52	"LIVING SINGLE"	5/10
...			

602   604   603   601

FIG. 6

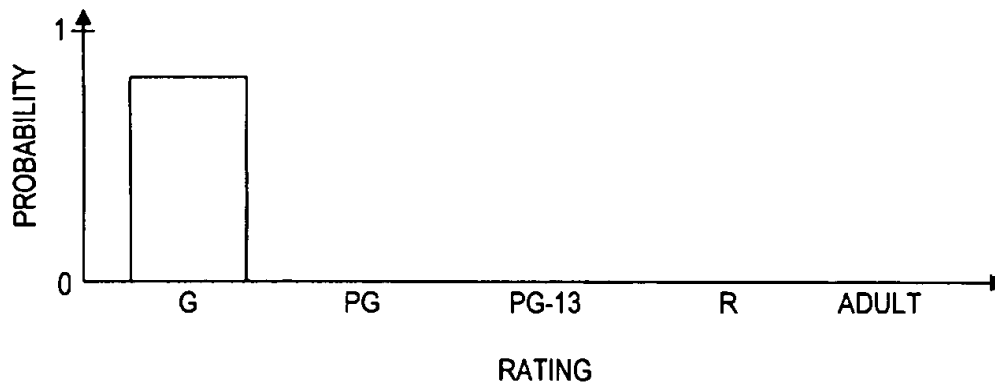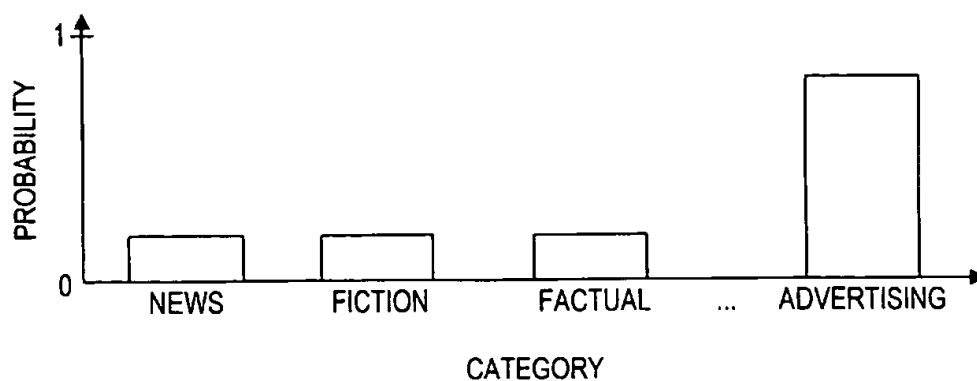| TIME OF DAY		MINUTES WATCHED	CHANNEL CHANGES	AVERAGE VOLUME
MORNING	(6AM-9AM)	61	2	5/10
MID-DAY	(9AM-3PM)	0	0	-
AFTERNOON	(3PM-6PM)	0	0	-
NIGHT	(6PM-10PM)	122	4	6/10
LATE NIGHT	(12AM-6AM)	0	0	-
TOTAL		183	6	5.7/10

700   702   704   706

FIG. 7
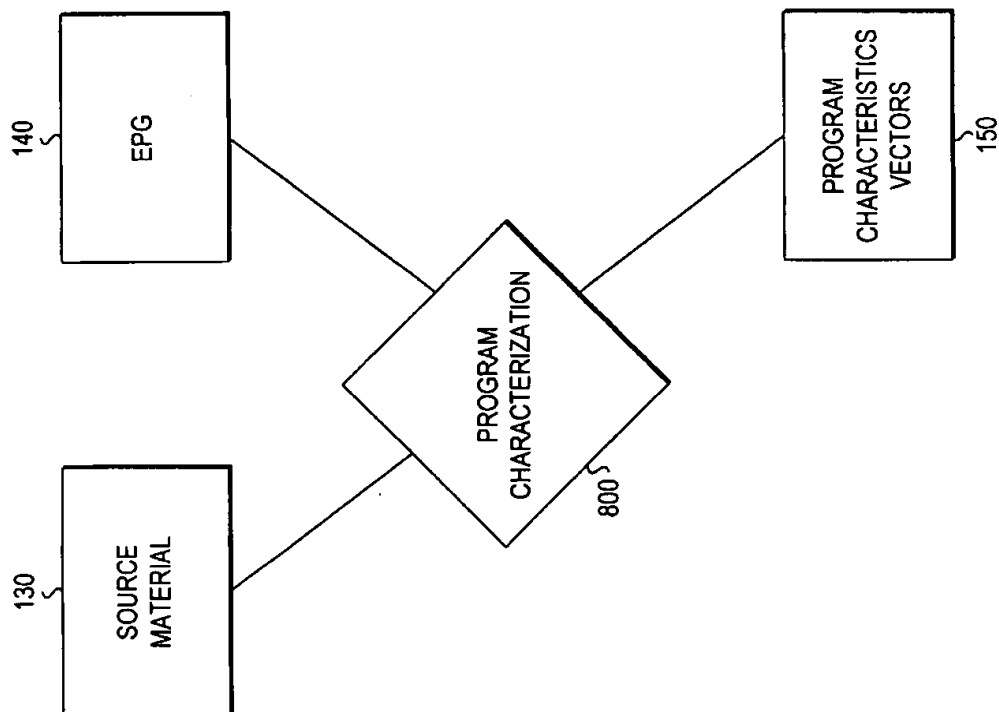
FIG. 8A
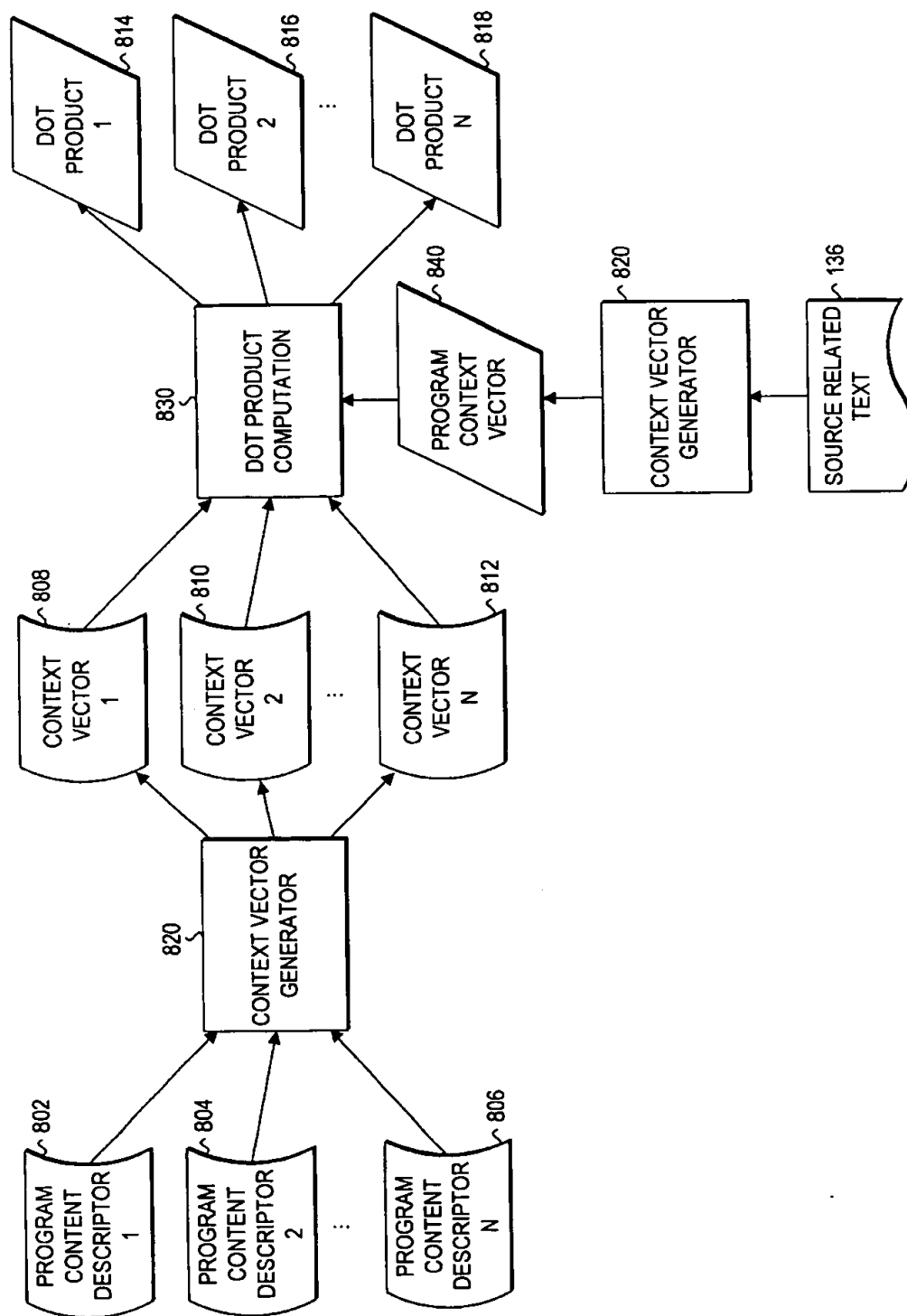
FIG. 8B
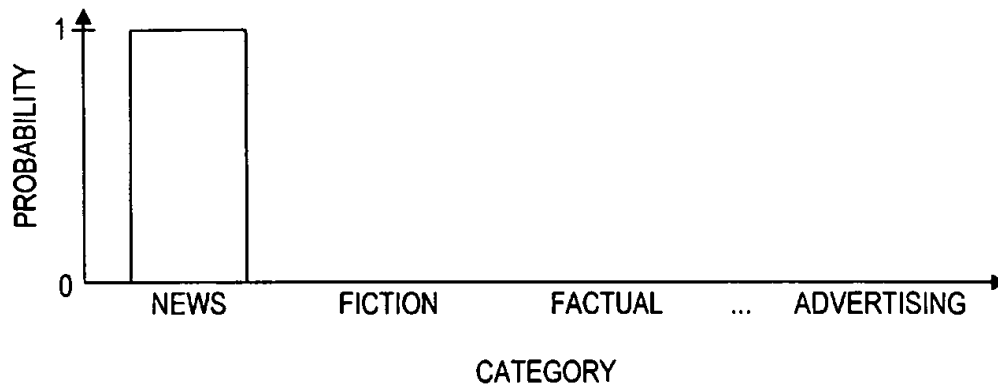
FIG. 9A
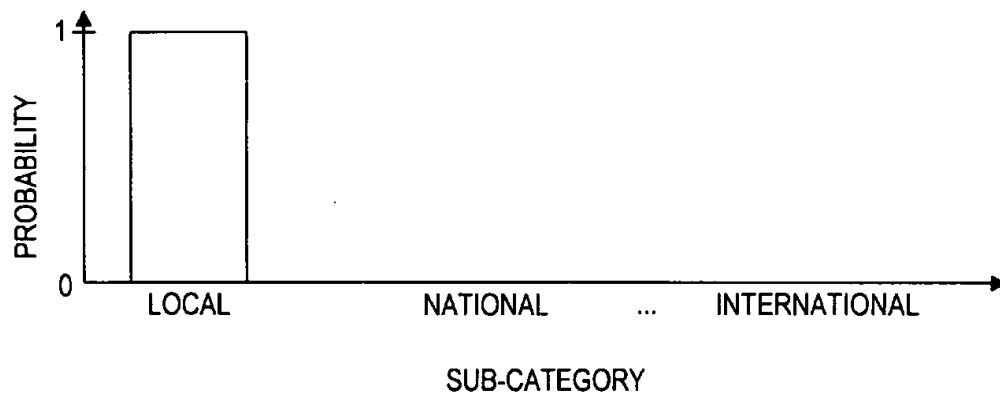
FIG. 9B

FIG. 9C

FIG. 9D

FIG. 9E

FIG. 9F

1030

YOUNG MEN CHANGE CHANNELS MORE OFTEN THAN YOUNG WOMEN.

1050

SOAP OPERA "DAYS OF OUR LIVES" WATCHED GENERALLY BY HOUSEWIVES.

1010

ANNUAL HOUSEHOLD INCOME	ZAPS ONCE EVERY
<$15,000	6 MINUTES, 15 SECONDS
$15-$24,999	4 MINUTES, 10 SECONDS
$25-$49,999	3 MINUTES, 27 SECONDS
$50-$74,999	3 MINUTES, 7 SECONDS
$75,000+	2 MINUTES, 42 SECONDS

HEURISTIC RULES
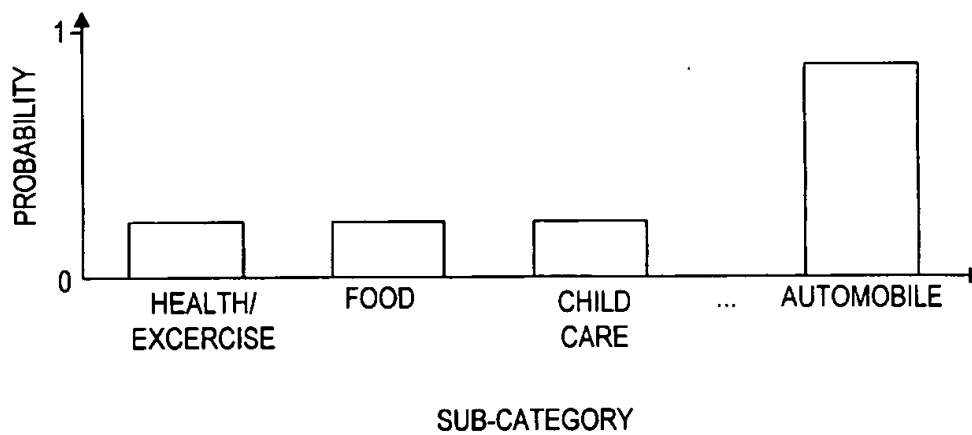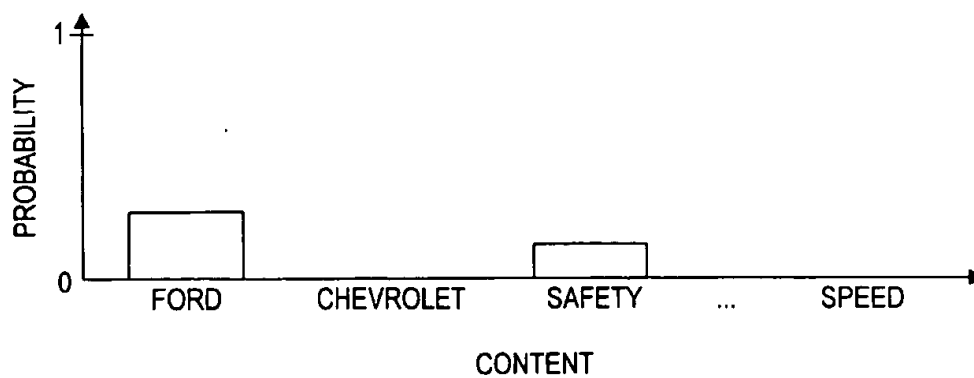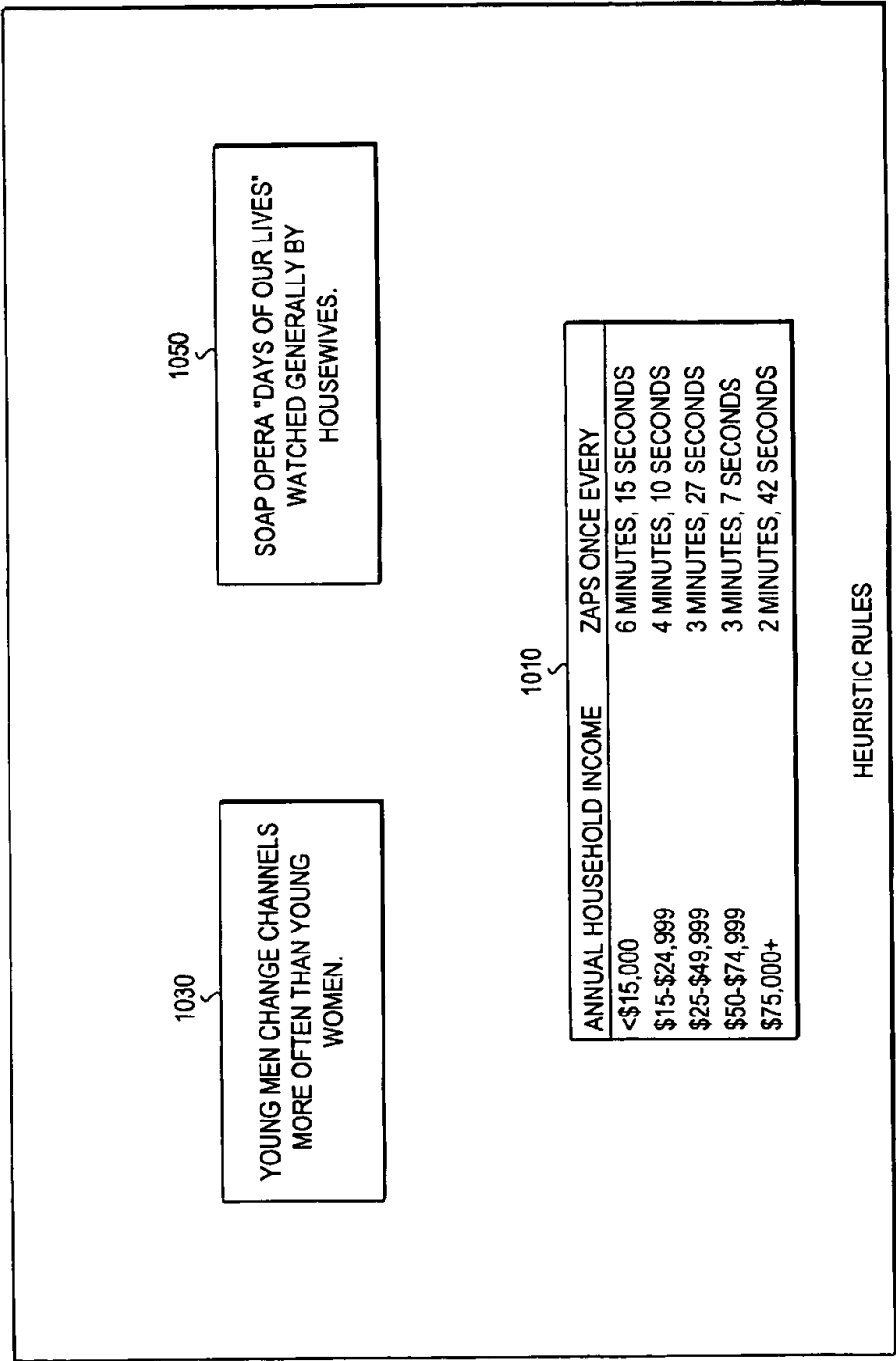
FIG. 10A

CATEGORIES	AGE			INCOME			SIZE			GENDER	
	0-10	10-18 ··	>70	0-20K	20-50K ··	50-100K	1	2 ··	>5	M	F
NEWS	0.1	0.1	0.4	0.2	0.3	0.4	0.5	0.3	0.1	0.3	0.7
FICTION	0.5	0.3	0.2	0.4	0.2	0.3	0.3	0.2	0.1	0.8	0.2
FACTUAL	0.2	0.2	0.3	0.1	0.4	0.2	0.2	0.2	0.4	0.4	0.6
...											
ADVERTISING	0.1	0.3	0.5	0.3	0.2	0.1	0.2	0.1	0.3	0.5	0.5

DEMOGRAPHIC GROUPS

FIG. 10B

150

PROGRAM CHARACTERISTIC VECTORS

160

HEURISTIC RULES

PROGRAM TARGET ANALYSIS

1100

PROGRAM DEMOGRAPHIC VECTORS

170
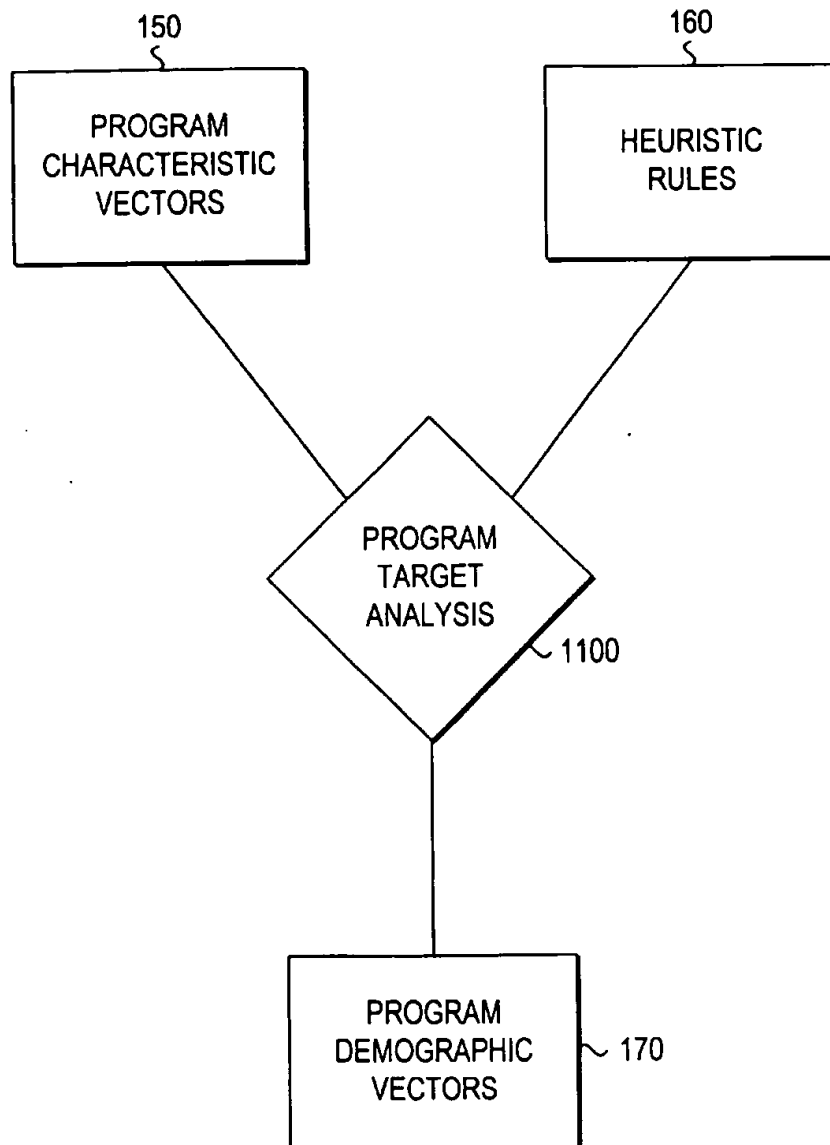
FIG. 11

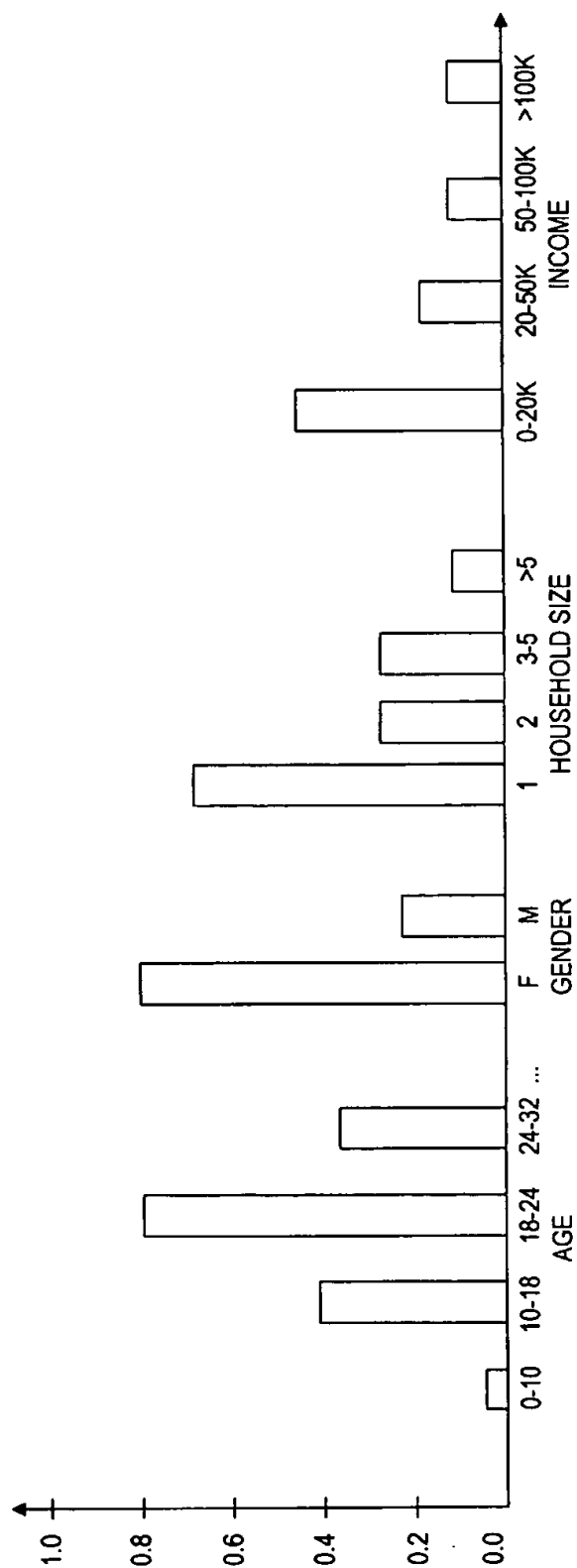*FIG. 12*
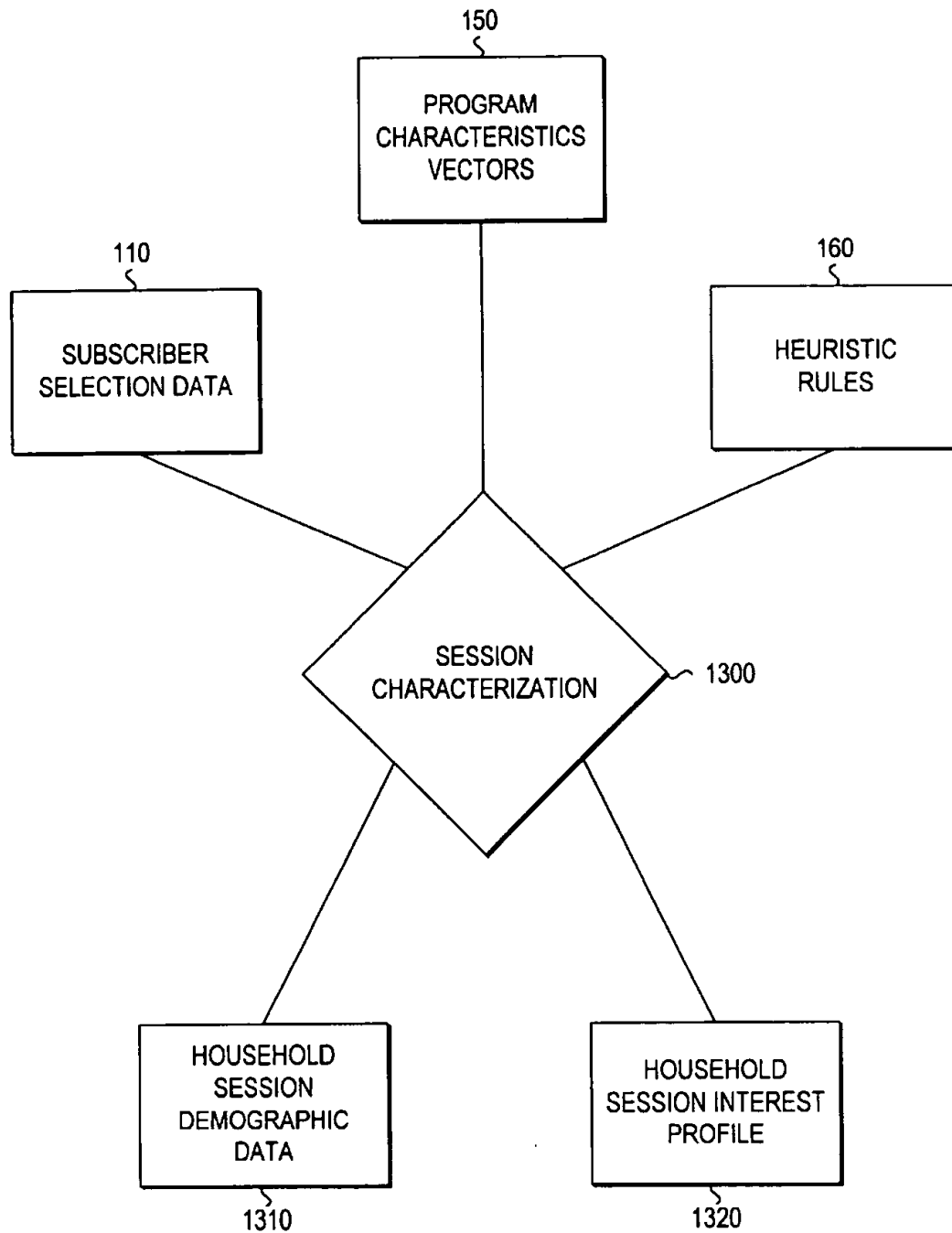
FIG. 13
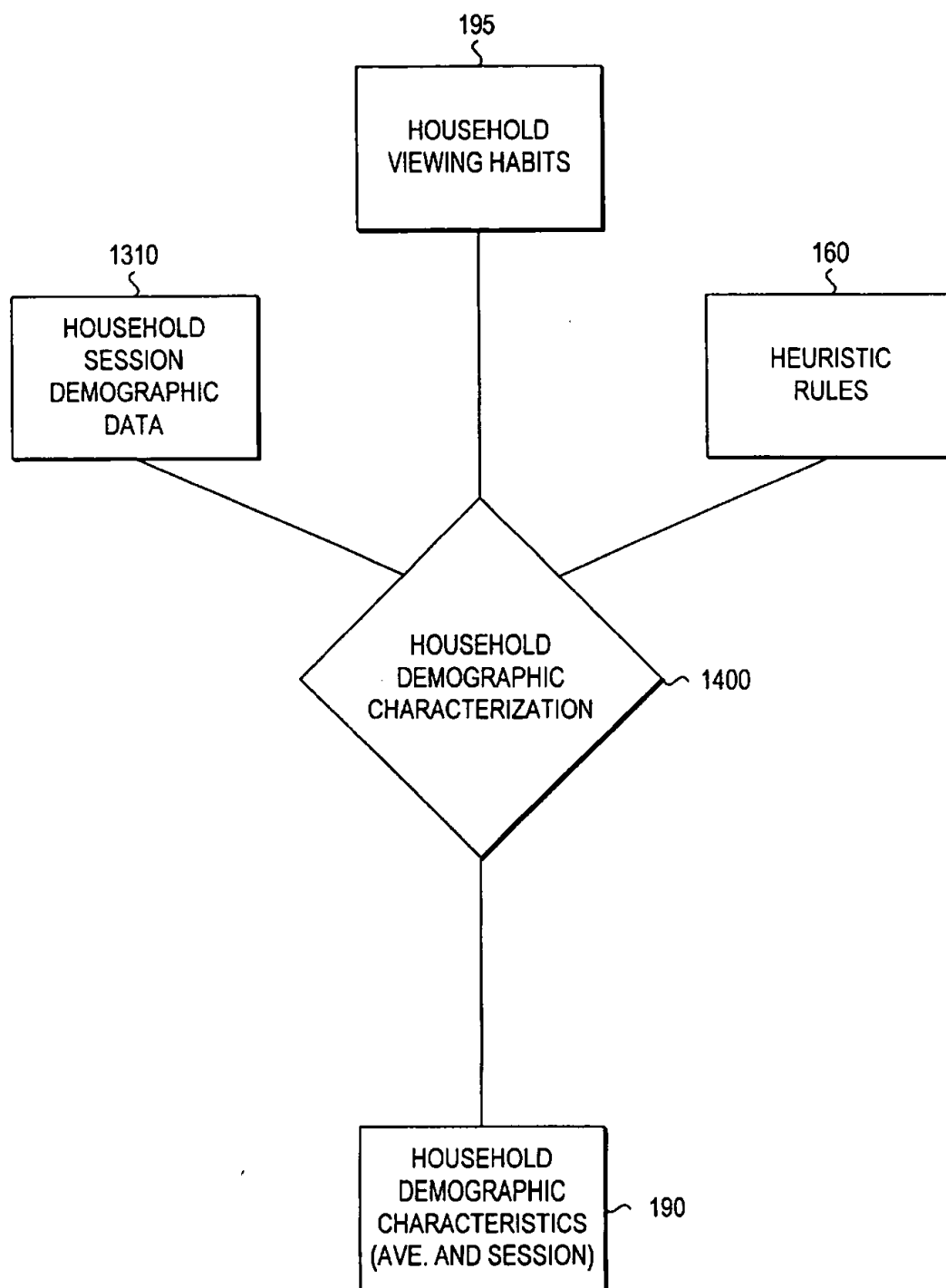
FIG. 14

HOUSEHOLD PARAMETER	AVERAGE VALUE	SESSION VALUE	UPDATE?
SIZE	2.6	3.0	YES
AGE	23.5	12	YES
SEX (FEMALE=1)	0.6	0.7	YES
INCOME ($0-$20K)	0.1	0.1	YES
INCOME ($20-$50K)	0.6	0.7	YES
INCOME ($50-$100K)	0.2	0.1	YES
INCOME (>$100K)	0.1	0.1	YES
ZIP CODE			NO
TELEPHONE NUMBER			NO

1501   1505   1503   1507

FIG. 15
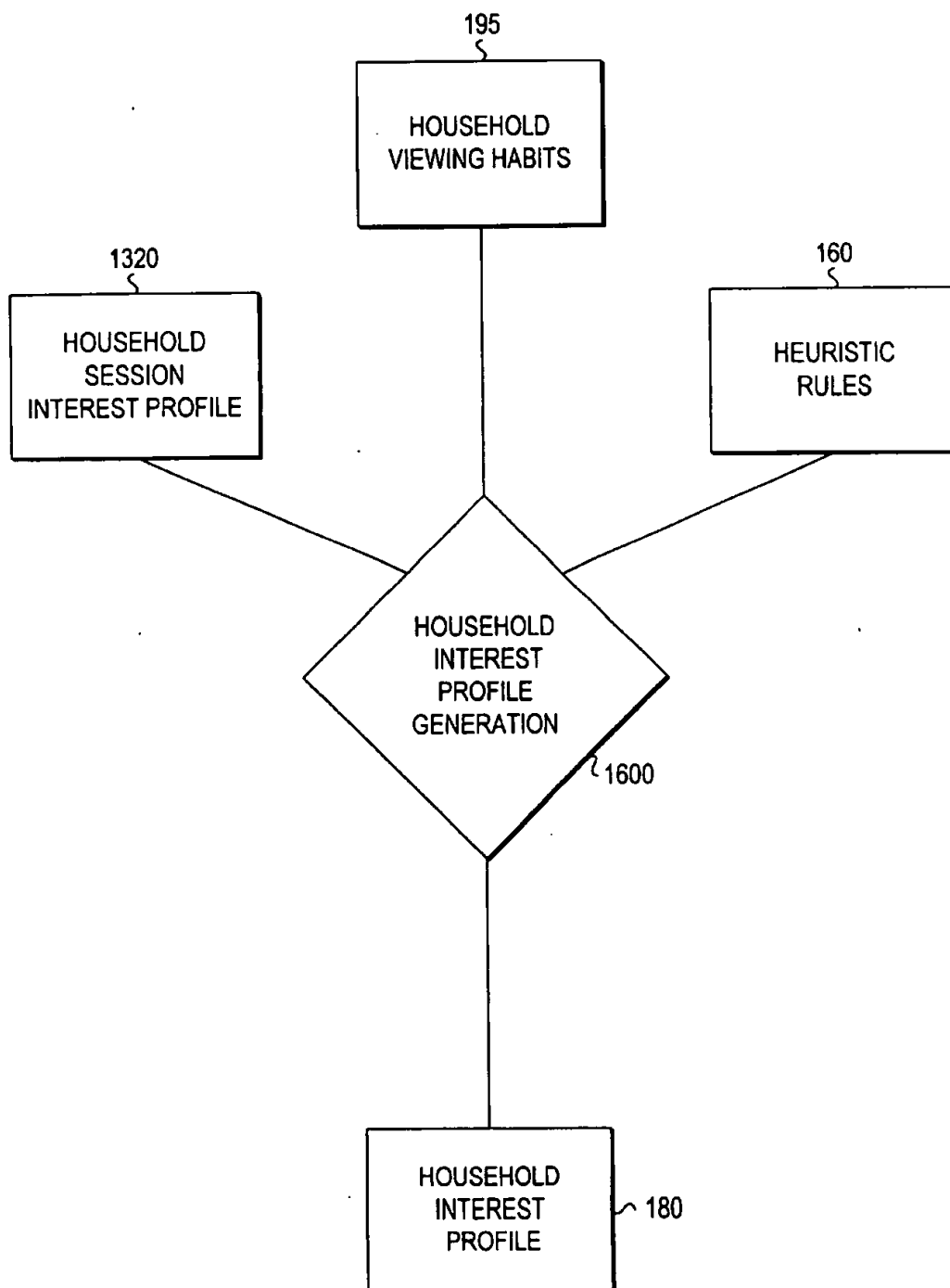
FIG. 16

	HOUSEHOLD INTEREST	AVERAGE VALUE	SESSION VALUE
	1701	*1703*	*1705*
PROGRAMMING	DRAMA	0.1	0.20
	ROMANCE	0.1	0.20
	ACTION	0.6	0.25
	SITCOM	0.2	0.30
	⋮		
	SPORTS	0	0.05
PRODUCTS	HEALTH/EXCERCISE	0.6	0.2
	FOOD	0.3	0.4
	CHILD RELATED	0.0	0.1
	TOYS	0.0	0.1
	⋮		
	AUTOMOBILE	0.1	0.2

1709 (PROGRAMMING)
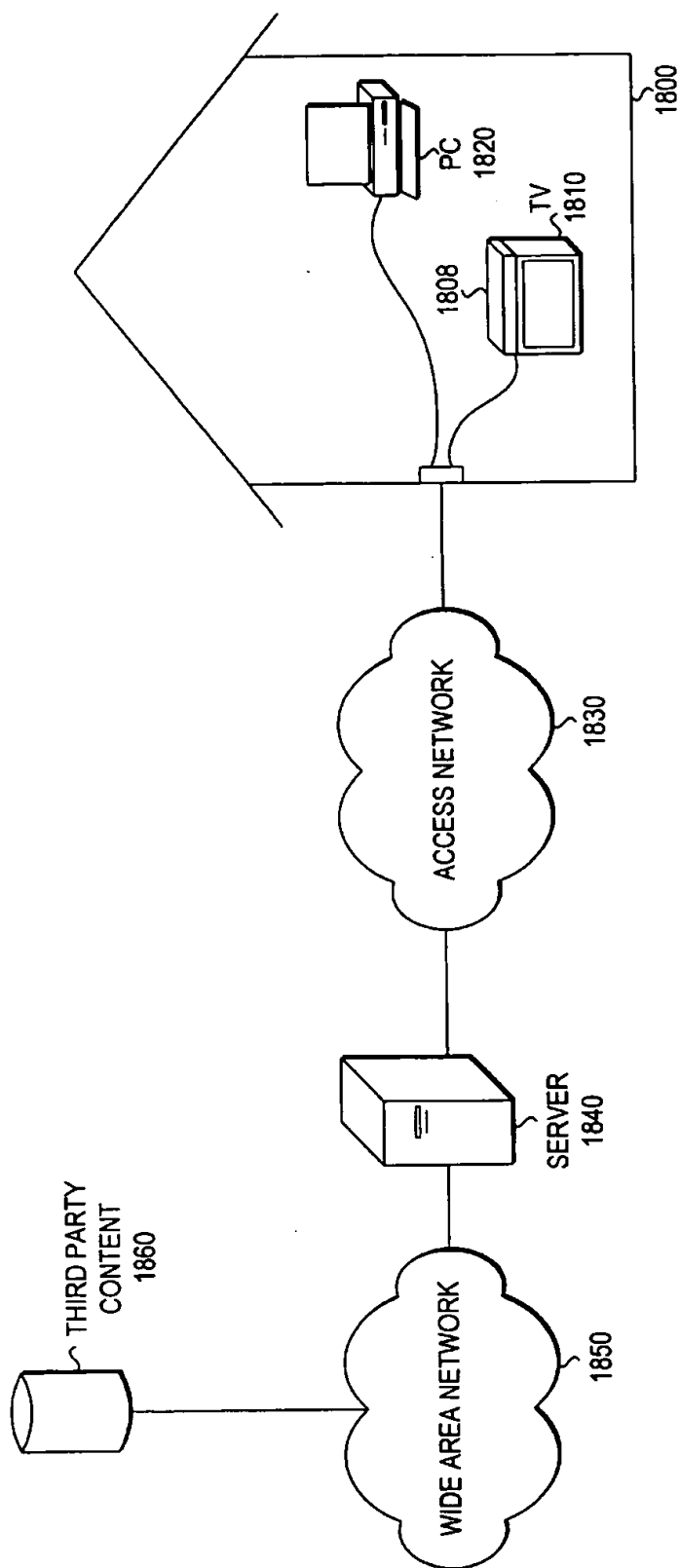1707 (PRODUCTS)

FIG. 17

FIG. 18

CLIENT-SERVER BASED SUBSCRIBER CHARACTERIZATION SYSTEM

BACKGROUND OF THE INVENTION

Subscribers face an increasingly large number of choices for entertainment programming, which is delivered over networks such as cable TV systems, over-the-air broadcast systems, and switched digital access systems which use telephone company twisted wire pairs for the deliver of signals.

Cable television service providers have typically provided one-way broadcast services but now offer high-speed data services and can combine traditional analog broadcasts with digital broadcasts and access to Internet web sites. Telephone companies can offer digital data and video programming on a switched basis over digital subscriber line technology. Although the subscriber may only be presented with one channel at a time, channel change requests are instantaneously transmitted to centralized switching equipment and the subscriber can access the programming in a broadcast-like manner. Internet Service Providers (ISPs) offer Internet access and can offer access to text, audio, and video programming which can also be delivered in a broadcast-like manner in which the subscriber selects "channels" containing programming of interest. Such channels may be offered as part of a video programming service or within a data service and can be presented within an Internet browser.

Along with the multitude of programming choices which the subscriber faces, subscribers are subject to advertisements, which in many cases subsidize or pay for the entire cost of the programming. While advertisements are sometimes beneficial to subscribers and deliver desired information regarding specific products or service, consumers generally vie advertising as a "necessary evil" for broadcast type entertainment.

In order to deliver more targeted programming and advertising to subscribers, it is necessary to understand their likes and dislikes to a greater extent than is presently done today. Systems which identify subscriber preferences based on their purchases and responses to questionnaires allow for the targeted marketing of literature in the mail, but do not in any sense allow for the rapid and precise delivery of programming and advertising which is known to have a high probability of acceptance to the subscriber. In order to determine which programming or advertising is appropriate for the subscriber, knowledge of that subscriber and the subscriber product and programming preferences is required.

Specific information regarding a subscriber's viewing habits or the Internet web sites they have accessed can be stored for analysis, but such records are considered private and subscribers are not generally willing to have such information leave their control. Although there are regulatory models which permit the collection of such data on a "notice and consent" basis, there is a general tendency towards legal rules which prohibit such raw data to be collected.

With the migration of services from a broadcast based model to a client-server based model in which subscribers make individualized request for programming to an Internet access provider or content provider, there is opportunity to monitor the subscriber viewing characteristics to better provide them with programming and advertising which will be of interest to them. A server may act as a proxy for the subscriber requests and thus be able to monitor what a

subscriber has requested and is viewing. Since subscribers may not want this raw data to be utilized, there is a need for a system which can process this information and generate statistically relevant subscriber profiles. These profiles should be accessible to others on the network who may wish to determine if their programming or advertisements are suitable for the subscriber.

For the foregoing reasons, there is a need for a subscriber characterization system which can generate and store subscriber characteristics which reflect the probable demographics and preferences of the subscriber and household.

SUMMARY OF THE INVENTION

The present invention includes a system for characterizing subscribers watching video or multimedia programming based on monitoring the requests made by the subscriber for programming to a server which contains the content or which requests the content from a third party. The server side of the network is able to monitor the subscriber's detailed selection choices including the time duration of their viewing, the volume the programming is listened at, and the program selection.

The server side collects text information about that programming to determine what type of programming the subscriber is most interested in. In addition the system can generate a demographic description of the subscriber or household which describes the probable age, income, gender and other demographics. The resulting characterization includes probabilistic determinations of what other programming or products the subscriber/household will be interested in.

In a preferred embodiment the textual information which describes the programming is obtained by context mining of text associated with the programming. The associated text can be from the closed-captioning data associated with the programming, an electronic program guide, or from text files associated with or part of the programming itself.

The system can provide both session measurements which correspond to a profile obtained over a viewing session, or an average profile which corresponds to data obtained over multiple viewing sessions.

The present invention also encompasses the use of heuristic rules in logical form or expressed as conditional probabilities to aid in forming a subscriber profile. The heuristic rules in logical form allow the system to apply generalizations which have been learned from external studies to obtain a characterization of the subscriber. In the case of conditional probabilities, determinations of the probable content of a program can be applied in a mathematical step to a matrix of conditional probabilities to obtain probabilistic subscriber profiles indicating program and product likes and dislikes as well for determining probabilistic demographic data.

One advantage of the present invention is that it allows consumers the possibility to permit access to probabilistic information regarding their household demographics and programming/product preferences, without revealing their specific viewing history. Subscribers may elect to permit access to this information in order to receive advertising which is more targeted to their likes/dislikes. Similarly, a subscriber may wish to sell access to this statistical data in order to receive revenue or receive a discount on a product or a service.

Another advantage of the present invention is that the resulting probabilistic information can be stored locally and controlled by the subscriber, or can be transferred to a third

party which can provide access to the subscriber characterization. The information can also be encrypted to prevent unauthorized access in which case only the subscriber of someone authorized by the subscriber can access the data.

These and other features and objects of the invention will be more fully understood from the following detailed description of the preferred embodiments which should be read in light of the accompanying drawings.

BRIEF DESCRIPTION OF THE DRAWINGS

The accompanying drawings, which are incorporated in and form a part of the specification, illustrate the embodiments of the present invention and, together with the description serve to explain the principles of the invention.

In the drawings:

FIG. 1 shows a context diagram for a subscriber characterization system.

FIG. 2 illustrates a block diagram for a realization of a subscriber monitoring system for receiving video signals;

FIG. 3 illustrates a block diagram of a channel processor;

FIG. 4 illustrates a block diagram of a computer for a realization of the subscriber monitoring system;

FIG. 5 illustrates a channel sequence and volume over a twenty-four (24) hour period;

FIG. 6 illustrates a time of day detailed record;

FIG. 7 illustrates a household viewing habits statistical table;

FIG. 8A illustrates an entity-relationship diagram for the generation of program characteristics vectors;

FIG. 8B illustrates a flowchart for program characterization;

FIG. 9A illustrates a deterministic program category vector;

FIG. 9B illustrates a deterministic program sub-category vector;

FIG. 9C illustrates a deterministic program rating vector;

FIG. 9D illustrates a probabilistic program category vector;

FIG. 9E illustrates a probabilistic program sub-category vector;

FIG. 9F illustrates a probabilistic program content vector;

FIG. 10A illustrates a set of logical heuristic rules;

FIG. 10B illustrates a set of heuristic rules expressed in terms of conditional probabilities;

FIG. 11 illustrates an entity-relationship diagram for the generation of program demographic vectors;

FIG. 12 illustrates a program demographic vector;

FIG. 13 illustrates an entity-relationship diagram for the generation of household session demographic data and household session interest profiles;

FIG. 14 illustrates an entity-relationship diagram for the generation of average and session household demographic characteristics;

FIG. 15 illustrates average and session household demographic data;

FIG. 16 illustrates an entity-relationship diagram for generation of a household interest profile;

FIG. 17 illustrates household interest profile including programming and product profiles; and

FIG. 18 illustrates a client-server architecture, realizing the present invention.

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

In describing a preferred embodiment of the invention illustrated in the drawings, specific terminology will be used

for the sake of clarity. However, the invention is not intended to be limited to the specific terms so selected, and it is to be understood that each specific term includes all technical equivalents which operate in a similar manner to accomplish a similar purpose.

With reference to the drawings, in general, and FIGS. 1 through 18 in particular, the apparatus of the present invention is disclosed.

The present invention is directed at an apparatus for generating a subscriber profile which contains useful information regarding the subscriber likes and dislikes. Such a profile is useful for systems which provide targeted programming or advertisements to the subscriber, and allow material (programs or advertisements) to be directed at subscribers who will have a high probability of liking the program or a high degree of interest in purchasing the product.

Since there are typically multiple individuals in a household, the subscriber characterization may not be a characterization of an individual subscriber but may instead be a household average. When used herein, the term subscriber refers both to an individual subscriber as well as the average characteristics of a household of multiple subscribers.

In the present system the programming viewed by the subscriber, both entertainment and advertisement, can be studied and processed by the subscriber characterization system to determine the program characteristics. This determination of the program characteristics is referred to as a program characteristics vector. The vector may be a truly one-dimensional vector, but can also be represented as an n dimensional matrix which can be decomposed into vectors.

The subscriber profile vector represents a profile of the subscriber (or the household of subscribers) and can be in the form of a demographic profile (average or session) or a program or product preference vector. The program and product preference vectors are considered to be part of a household interest profile which can be thought of as an n dimensional matrix representing probabilistic measurements of subscriber interests.

In the case that the subscriber profile vector is a demographic profile, the subscriber profile vector indicates a probabilistic measure of the age of the subscriber or average age of the viewers in the household, sex of the subscriber, income range of the subscriber or household, and other such demographic data. Such information comprises household demographic characteristics and is composed of both average and session values. Extracting a single set of values from the household demographic characteristics can correspond to a subscriber profile vector.

The household interest profile can contain both programming and product profiles, with programming profiles corresponding to probabilistic determinations of what programming the subscriber (household) is likely to be interested in, and product profiles corresponding to what products the subscriber (household) is likely to be interested in. These profiles contain both an average value and a session value, the average value being a time average of data, where the averaging period may be several days, weeks, months, or the time between resets of unit.

Since a viewing session is likely to be dominated by a particular viewer, the session values may, in some circumstances, correspond most closely to the subscriber values, while the average values may, in some circumstances, correspond most closely to the household values.

5

FIG. 1 depicts the context diagram of a preferred embodiment of a Subscriber Characterization System (SCS) 100. A context diagram, in combination with entity-relationship diagrams, provide a basis from which one skilled in the art can realize the present invention. The present invention can be realized in a number of programming languages including C, C++, Perl, and Java, although the scope of the invention is not limited by the choice of a particular programming language or tool. Object oriented languages have several advantages in terms of construction of the software used to realize the present invention, although the present invention can be realized in procedural or other types of programming languages known to those skilled in the art.

In generating a subscriber profile, the SCS 100 receives from a user 120 commands in the form of a volume control signal 124 or program selection data 122 which can be in the form of a channel change but may also be an address request which requests the delivery of programming from a network address. A record signal 126 indicates that the programming or the address of the programming is being recorded by the user. The record signal 126 can also be a printing command, a tape recording command, a bookmark command or any other command intended to store the program being viewed, or program address, for later use.

The material being viewed by the user 12 is referred to as source material 130. The source material 130, as defined herein, is the content that a subscriber selects and may consist of analog video, Motion Picture Expert Group (MPEG) digital video source material, other digital or analog material, Hypertext Markup Language (HTML) or other type of multimedia source material. The subscriber characterization system 100 can access the source material 130 received by the user 120 using a start signal 132 and a stop signal 134, which control the transfer of source related text 136 which can be analyzed as described herein.

In a preferred embodiment, the source related text 136 can be extracted from the source material 130 and stored in memory. The source related text 136, as defined herein, includes source related textual information including descriptive fields which are related to the source material 130, or text which is part of the source material 130 itself. The source related text 136 can be derived from a number of sources including but not limited to closed captioning information, Electronic Program Guide (EPG) material, and text information in the source itself (e.g. text in HTML files).

Electronic Program Guide (EPG) 140 contains information related to the source material 130 which is useful to the user 120. The EPG 140 is typically a navigational tool which contains source related information including but not limited to the programming category, program description, rating, actors, and duration. The structure and content of EPG data is described in detail in U.S. Pat. No. 5,596,373 assigned to Sony Corporation and Sony Electronics which is herein incorporated by reference. As shown in FIG. 1, the EPG 140 can be accessed by the SCS 100 by a request EPG data signal 142 which results in the return of a category 144, a sub-category 146, and a program description 148.

In one embodiment of the present invention, EPG data is accessed and program information such as the category 144, the sub-category 146, and the program description 148 are stored in memory.

In another embodiment of the present invention, the source related text 136 is the closed captioning text embedded in the analog or digital video signal. Such closed captioning text can be stored in memory for processing to extract the program characteristic vectors 150.

6

One of the functions of the SCS 100 is to generate the program characteristics vectors 150 which are comprised of program characteristics data 152, as illustrated in FIG. 1. The program characteristics data 152, which can be used to create the program characteristics vectors 150 both in vector and table form, are examples of source related information which represent characteristics of the source material. In a preferred embodiment, the program characteristics vectors 150 are lists of values which characterize the programming (source) material in according to the category 144, the sub-category 146, and the program description 148. The present invention may also be applied to advertisements, in which case program characteristics vectors contain, as an example, a product category, a product sub-category, and a brand name.

As illustrated in FIG. 1, the SCS 100 uses heuristic rules 160. The heuristic rules 160, as described herein, are composed of both logical heuristic rules as well as heuristic rules expressed in terms of conditional probabilities. The heuristic rules 160 can be accessed by the SCS 100 via a request rules signal 162 which results in the transfer of a copy of rules 164 and the SCS 100.

The SCS 100 forms program demographic vectors 170 from program demographics 172, as illustrated in FIG. 1. The program demographic vectors 170 also represent characteristics of source related information in the form of the intended or expected demographics of the audience for which the source material is intended.

Subscriber selection data 110 is obtained from the monitored activities of the user and in a preferred embodiment can be stored in a dedicated memory. In an alternate embodiment, the subscriber selection data 110 is stored in a storage disk. Information which is utilized to form the subscriber selection data 110 includes time 112, which corresponds to the time of an event, channel ID 114, program ID 116, volume level 118, channel change record 119, and program title 117. A detailed record of selection data is illustrated in FIG. 6.

In a preferred embodiment, a household viewing habits 195 illustrated in FIG. 1 is computed from the subscriber selection data 110. The SCS 100 transfers household viewing data 197 to form household viewing habits 195. The household viewing data 197 is derived from the subscriber selection data 110 by looking at viewing habits at a particular time of day over an extended period of time, usually several days or weeks, and making some generalizations regarding the viewing habits during that time period.

The program characteristics vector 150 is derived from the source related text 136 and/or from the EPG 140 by applying information retrieval techniques. The details of this process are discussed in accordance with FIG. 8.

The program characteristics vector 150 is used in combination with a set of the heuristic rules 160 to define a set of the program demographic vectors 170 illustrated in FIG. 1 describing the audience the program is intended for.

One output of the SCS 100 is a household profile including household demographic characteristics 190 and a household interest profile 180. The household demographic characteristics 190 resulting from the transfer of household demographic data 192, and the household interest profile 180, resulting from the transfer of household interests data 182. Both the household demographics characteristics 190 and the household interest profile 180 have a session value and an average value, as will be discussed herein.

The monitoring system depicted in FIG. 2 is responsible for monitoring the subscriber activities, and can be used to

7

realize the SCS 100. In a preferred embodiment, the monitoring system of FIG. 2 is located in a television set-top device or in the television itself. In an alternate embodiment, the monitoring system is part of a computer which receives programming from a network.

In an application of the system for television services, an input connector 220 accepts the video signal coming either from an antenna, cable television input, or other network. The video signal can be analog or Digital MPEG. Alternatively, the video source may be a video stream or other multimedia stream from a communications network including the Internet.

In the case of either analog or digital video, selected fields are defined to carry EPG data or closed captioning text. For analog video, the closed captioning text is embedded in the vertical blanking interval (VBI). As described in U.S. Pat. No. 5,579,005, assigned to Scientific-Atlanta, Inc., the EPG information can be carried in a dedicated channel or embedded in the VBI. For digital video, the closed captioning text is carried as video user bits in a user_data field. The EPG data is transmitted as ancillary data and is multiplexed at the transport layer with the audio and video data.

Referring to FIG. 2, a system control unit 200 receives commands from the user 120, decodes the command and forwards the command to the destined module. In a preferred embodiment, the commands are entered via a remote control to a remote receiver 205 or a set of selection buttons 207 available at the front panel of the system control unit 200. In an alternate embodiment, the commands are entered by the user 120 via a keyboard.

The system control unit 200 also contains a Central Processing Unit (CPU) 203 for processing and supervising all of the operations of the system control unit 200, a Read Only Memory (ROM) 202 containing the software and fixed data, a Random Access Memory (RAM) 204 for storing data. CPU 203, RAM 204, ROM 202, and I/O controller 201 are attached to a master bus 206. A power supply in a form of battery can also be included in the system control unit 200 for backup in case of power outage.

An input/output (I/O) controller 201 interfaces the system control unit 200 with external devices. In a preferred embodiment, the I/O controller 201 interfaces to the remote receiver 205 and a selection button such as the channel change button on a remote control. In an alternate embodiment, it can accept input from a keyboard or a mouse.

The program selection data 122 is forwarded to a channel processor 210. The channel processor 210 tunes to a selected channel and the media stream is decomposed into its basic components: the video stream, the audio stream, and the data stream. The video stream is directed to a video processor module 230 where it is decoded and further processed for display to the TV screen. The audio stream is directed to an audio processor 240 for decoding and output to the speakers.

The data stream can be EPG data, closed captioning text, Extended Data Service (EDS) information, a combination of these, or an alternate type of data. In the case of EDS the call sign, program name and other useful data are provided. In a preferred embodiment, the data stream is stored in a reserved location of the RAM 204. In an alternate embodiment, a magnetic disk is used for data storage. The system control unit 200 writes also in a dedicated memory, which in a preferred embodiment is the RAM 204, the selected channel, the time 112 of selection, the volume level 118 and the program ID 116 and the program title 117. Upon receiving the program selection data 122, the new selected channel is

8

directed to the channel processor 210 and the system control unit 200 writes to the dedicated memory the channel selection end time and the program title 117 at the time 112 of channel change. The system control unit 200 keeps track of the number of channel changes occurring during the viewing time via the channel change record 119. This data forms part of the subscriber selection data 110.

The volume control signal 124 is sent to the audio processor 240. In a preferred embodiment, the volume level 118 selected by the user 120 corresponds to the listening volume. In an alternate embodiment, the volume level 118 selected by the user 120 represents a volume level to another piece of equipment such as an audio system (home theatre system) or to the television itself. In such a case, the volume can be measured directly by a microphone or other audio sensing device which can monitor the volume at which the selected source material is beings listened.

A program change occurring while watching a selected channel is also logged by the system control unit 200. Monitoring the content of the program at the time of the program change can be done by reading the content of the EDS. The EDS contains information such as program title, which is transmitted via the VBI. A change on the program title field is detected by the monitoring system and logged as an event. In an alternate embodiment, an EPG is present and program information can be extracted from the EPG. In a preferred embodiment, the programming data received from the EDS or EPG permits distinguishing between entertainment programming and advertisements.

FIG. 3 shows the block diagram of the channel processor 210. In a preferred embodiment, the input connector 220 connects to a tuner 300 which tunes to the selected channel. A local oscillator can be used to heterodyne the signal to the IF signal. A demodulator 302 demodulates the received signal and the output is fed to an FEC decoder 304. The data stream received from the FEC decoder 304 is, in a preferred embodiment, in an MPEG format. In a preferred embodiment, system demultiplexer 306 separates out video and audio information for subsequent decompression and processing, as well as ancillary data which can contain program related information.

The data stream presented to the system demultiplexer 306 consists of packets of data including video, audio and ancillary data. The system demultiplexer 306 identifies each packet from the stream ID and directs the stream to the corresponding processor. The video data is directed to the video processor module 230 and the audio data is directed to the audio processor 240. The ancillary data can contain closed captioning text, emergency messages, program guide, or other useful information.

Closed captioning text is considered to be ancillary data and is thus contained in the video stream. The system demultiplexer 306 accesses the user data field of the video stream to extract the closed captioning text. The program guide, if present, is carried on data stream identified by a specific transport program identifier.

In an alternate embodiment, analog video can be used. For analog programming, ancillary data such as closed captioning text or EDS data are carried in a vertical blanking interval.

FIG. 4 shows the block diagram of a computer system for a realization of the subscriber monitoring system based on the reception of multimedia signals from a bi-directional network. A system bus 422 transports data amongst the CPU 203, the RAM 204, Read Only Memory—Basic Input Output System (ROM-BIOS) 406 and other components.

9

10

The CPU 203 accesses a hard drive 400 through a disk controller 402. The standard input/output devices are connected to the system bus 422 through the I/O controller 201. A keyboard is attached to the I/O controller 201 through a keyboard port 416 and the monitor is connected through a monitor port 418. The serial port device uses a serial port 420 to communicate with the I/O controller 201. Industry Standard Architecture (ISA) expansion slots 408 and Peripheral Component Interconnect (PCI) expansion slots 410 allow additional cards to be placed into the computer. In a preferred embodiment, a network card is available to interface a local area, wide area, or other network.

FIG. 5 illustrates a channel sequence and volume over a twenty-four (24) hour period. The Y-axis represents the status of the receiver in terms of on/off status and volume level. The X-axis represents the time of day. The channels viewed are represented by the windows 501–506, with a first channel 502 being watched followed by the viewing of a second channel 504, and a third channel 506 in the morning. In the evening a fourth channel 501 is watched, a fifth channel 503, and a sixth channel 505. A channel change is illustrated by a momentary transition to the "off" status and a volume change is represented by a change of level on the Y-axis.

A detailed record of the subscriber selection data 110 is illustrated in FIG. 6 in a table format. A time column 602 contains the starting time of every event occurring during the viewing time. A Channel ID column 604 lists the channels viewed or visited during that period. A program title column 603 contains the titles of all programs viewed. A volume column 601 contains the volume level 118 at the time 112 of viewing a selected channel.

A representative statistical record corresponding to the household viewing habits 195 is illustrated in FIG. 7. In a preferred embodiment, a time of day column 700 is organized in period of time including morning, mid-day, afternoon, night, and late night. In an alternate embodiment, smaller time periods are used. A minutes watched column 702 lists, for each period of time, the time in minutes in which the SCS 100 recorded delivery of programming. The number of channel changes during that period and the average volume are also included in that table in a channel changes column 704 and an average volume column 706 respectively. The last row of the statistical record contains the totals for the items listed in the minutes watched column 702, the channel changes column 704 and the average volume 706.

FIG. 8A illustrates an entity-relationship diagram for the generation of the program characteristics vector 150. The context vector generation and retrieval technique described in U.S. Pat. No. 5,619,709, which is incorporated herein by reference, can be applied for the generation of the program characteristics vectors 150. Other techniques are well known by those skilled in the art.

Referring to FIG. 8A, the source material 130 or the EPG 140 are passed through a program characterization process 800 to generate the program characteristics vectors 150. The program characterization process 800 is described in accordance with FIG. 8B. Program content descriptors including a first program content descriptor 802, a second program content descriptor 804 and an nth program content descriptor 806, each classified in terms of the category 144, the sub-category 146, and other divisions as identified in the industry accepted program classification system, are presented to a context vector generator 820. As an example, the program content descriptor can be text representative of the expected content of material found in the particular program category 144. In this example, the program content descriptors 802, 804 and 806 would contain text representative of what would be found in programs in the news, fiction, and advertising categories respectively. The context vector generator 820 generates context vectors for that set of sample texts resulting in a .first summary context vector 808, a second summary context vector 810, and an nth summary context vector 812. In the example given, the summary context vectors 808, 810, and 812 correspond to the categories of news, fiction and advertising respectively. The summary vectors are stored in a local data storage system.

Referring to FIG. 8B, a sample of the source related text 136 which is associated with the new program to be classified is passed to the context vector generator 820 which generates a program context vector 840 for that program. The source related text 136 can be either the source material 130, the EPG 140, or other text associated with the source material. A comparison is made between the actual program context vectors and the stored program content context vectors by computing, in a dot product computation process 830, the dot product of the first summary context vector 808 with the program context vector 840 to produce a first dot product 814. Similar operations are performed to produce second dot product 816 and nth dot product 818.

The values contained in the dot products 814, 816 and 818, while not probabilistic in nature, can be expressed in probabilistic terms using a simple transformation in which the result represents a confidence level of assigning the corresponding content to that program. The transformed values add up to one. The dot products can be used to classify a program, or form a weighted sum of classifications which results in the program characteristics vectors 150. In the example given, if the source related text 136 was from an advertisement, the nth dot product 818 would have a high value, indicating that the advertising category was the most appropriate category, and assigning a high probability value to that category. If the dot products corresponding to the other categories were significantly higher than zero, those categories would be assigned a value, with the result being the program characteristics vectors 150 as shown in FIG. 9D.

For the sub-categories, probabilities obtained from the content pertaining to the same sub-category 146 are summed to form the probability for the new program being in that sub-category 146. At the sub-category level, the same method is applied to compute the probability of a program being from the given category 144. The three levels of the program classification system; the category 144, the sub-category 146 and the content, are used by the program characterization process 800 to form the program characteristics vectors 150 which are depicted in FIGS. 9D–9F.

The program characteristics vectors 150 in general are represented in FIGS. 9A through 9F. FIGS. 9A, 9B and 9C are an example of deterministic program vectors. This set of vectors is generated when the program characteristics are well defined, as can occur when the source related text 136 or the EPG 140 contains specific fields identifying the category 144 and the sub-category 146. A program rating can also provided by the EPG 140.

In the case that these characteristics are not specified, a statistical set of vectors is generated from the process described in accordance with FIG. 8. FIG. 9D shows the probability that a program being watched is from the given category 144. The categories are listed in the X-axis. The sub-category 146 is also expressed in terms of probability.

This is shown in FIG. 9E. The content component of this set of vectors is a third possible level of the program classification, and is illustrated in FIG. 9F.

FIG. 10A illustrates sets of logical heuristics rules which form part of the heuristic rules 160. In a preferred embodiment, logical heuristic rules are obtained from sociological or psychological studies. Two types of rules are illustrated in FIG. 10A. The first type links an individual's viewing characteristics to demographic characteristics such as gender, age, and income level. A channel change rate rule 1030 attempts to determine gender based on channel change rate. An income related channel change rate rule 1010 attempts to link channel change rates to income brackets. A second type of rules links particular programs to particular audience, as illustrated by a gender determining rule 1050 which links the program category 144/sub-category 146 with a gender. The result of the application of the logical heuristic rules illustrated in FIG. 10A are probabilistic determinations of factors including gender, age, and income level. Although a specific set of logical heuristic rules has been used as an example, a wide number of types of logical heuristic rules can be used to realize the present invention. In addition, these rules can be changed based on learning within the system or based on external studies which provide more accurate rules.

FIG. 10B illustrates a set of the heuristic rules 160 expressed in terms of conditional probabilities. In the example shown in FIG. 10B, the category 144 has associated with it conditional probabilities for demographic factors such as age, income, family size and gender composition. The category 144 has associated with it conditional probabilities that represent probability that the viewing group is within a certain age group dependent on the probability that they are viewing a program in that category 144.

FIG. 11 illustrates an entity-relationship diagram for the generation of the program demographic vectors 170. In a preferred embodiment, the heuristic rules 160 are applied along with the program characteristic vectors 150 in a program target analysis process 1100 to form the program demographic vectors 170. The program characteristic vectors 150 indicate a particular aspect of a program, such as its violence level. The heuristic rules 160 indicate that a particular demographic group has a preference for that program. As an example, it may be the case that young males have a higher preference for violent programs than other sectors of the population. Thus, a program which has the program characteristic vectors 150 indicating a high probability of having violent content, when combined with the heuristic rules 160 indicating that "young males like violent programs," will result, through the program target analysis process 1100, in the program demographic vectors 170 which indicate that there is a high probability that the program is being watched by a young male.

The program target analysis process 1100 can be realized using software programmed in a variety of languages which processes mathematically the heuristic rules 160 to derive the program demographic vectors 170. The table representation of the heuristic rules 160 illustrated in FIG. 10B expresses the probability that the individual or household is from a specific demographic group based on a program with a particular category 144. This can be expressed, using probability terms as follow "the probability that the individuals are in a given demographic group conditional to the program being in a given category". Referring to FIG. 9D, the probability that the group has certain demographic characteristics based on the program being in a specific category is illustrated.

Expressing the probability that a program is destined to a specific demographic group can be determined by applying Bayes rule. This probability is the sum of the conditional probabilities that the demographic group likes the program, conditional to the category 144 weighted by the probability that the program is from that category 144. In a preferred embodiment, the program target analysis can calculate the program demographic vectors by application of logical heuristic rules, as illustrated in FIG. 10A, and by application of heuristic rules expressed as conditional probabilities as shown in FIG. 10B. Logical heuristic rules can be applied using logical programming and fuzzy logic using techniques well understood by those skilled in the art, and are discussed in the text by S. V. Kartalopoulos entitled "Understanding Neural Networks and Fuzzy Logic" which is incorporated herein by reference.

Conditional probabilities can be applied by simple mathematical operations multiplying program context vectors by matrices of conditional probabilities. By performing this process over all the demographic groups, the program target analysis process 1100 can measure how likely a program is to be of interest to each demographic group. Those probabilities values form the program demographic vector 170 represented in FIG. 12.

As an example, the heuristic rules expressed as conditional probabilities shown in FIG. 10B are used as part of a matrix multiplication in which the program characteristics vector 150 of dimension N, such as those shown in FIGS. 9A–9F is multiplied by an N×M matrix of heuristic rules expressed as conditional probabilities, such as that shown in FIG. 10B. The resulting vector of dimension M is a weighted average of the conditional probabilities for each category and represents the household demographic characteristics 190. Similar processing can be performed at the sub-category and content levels.

FIG. 12 illustrates an example of the program demographic vector 170, and shows the extent to which a particular program is destined to a particular audience. This is measured in terms of probability as depicted in FIG. 12. The Y-axis is the probability of appealing to the demographic group identified on the X-axis.

FIG. 13 illustrates an entity-relationship diagram for the generation of household session demographic data 1310 and household session interest profile 1320. In a preferred embodiment, the subscriber selection data 110 is used along with the program characteristics vectors 150 in a session characterization process 1300 to generate the household session interest profile 1320. The subscriber selection data 110 indicates what the subscriber is watching, for how long and at what volume they are watching the program.

In a preferred embodiment, the session characterization process 1300 forms a weighted average of the program characteristics vectors 150 in which the time duration the program is watched is normalized to the session time (typically defined as the time from which the unit was turned on to the present). The program characteristics vectors 150 are multiplied by the normalized time duration (which is less than one unless only one program has been viewed) and summed with the previous value. Time duration data, along with other subscriber viewing information, is available from the subscriber selection data 110. The resulting weighted average of program characteristics vectors forms the household session interest profile 1320, with each program contributing to the household session interest profile 1320 according to how long it was watched. The household session interest profile 1320 is normalized to produce proba-

bilistic values of the household programming interests during that session.

In an alternate embodiment, the heuristic rules 160 are applied to both the subscriber selection data 110 and the program characteristics vectors 150 to generate the household session demographic data 1310 and the household session interest profile 1320. In this embodiment, weighted averages of the program characteristics vectors 150 are formed based on the subscriber selection data 110, and the heuristic rules 160 are applied. In the case of logical heuristic rules as shown in FIG. 10A, logical programming can be applied to make determinations regarding the household session demographic data 1310 and the household session interest profile 1320. In the case of heuristic rules in the form of conditional probabilities such as those illustrated in FIG. 10B, a dot product of the time averaged values of the program characteristics vectors can be taken with the appropriate matrix of heuristic rules to generate both the household session demographic data 1310 and the household session interest profile 1320.

Volume control measurements which form part of the subscriber selection data 110 can also be applied in the session characterization process 1300 to form a household session interest profile 1320. This can be accomplished by using normalized volume measurements in a weighted average manner similar to how time duration is used. Thus, muting a show results in a zero value for volume, and the program characteristics vector 150 for this show will not be averaged into the household session interest profile 1320.

FIG. 14 illustrates an entity-relationship diagram for the generation of average household demographic characteristics and session household demographic characteristics 190. A household demographic characterization process 1400 generates the household demographic characteristics 190 represented in table format in FIG. 15. The household demographic characterization process 1400 uses the household viewing habit 195 in combination with the heuristic rules 160 to determine demographic data. For example, a household with a number of minutes watched of zero during the day may indicate a household with two working adults. Both logical heuristic rules as well as rules based on conditional probabilities can be applied to the household viewing habits 195 to obtain the household demographics characteristics 190.

The household viewing habits 195 is also used by the system to detect out-of-habits events. For example, if a household with a zero value for the minutes watched column 702 at late night presents a session value at that time via the household session demographic data 1310, this session will be characterized as an out-of-habits event and the system can exclude such data from the average if it is highly probable that the demographics for that session are greatly different than the average demographics for the household. Nevertheless, the results of the application of the household demographic characterization process 1400 to the household session demographic data 1310 can result in valuable session demographic data, even if such data is not added to the average demographic characterization of the household.

FIG. 15 illustrates the average and session household demographic characteristics. A household demographic parameters column 1501 is followed by an average value column 1505, a session value column 1503, and an update column 1507. The average value column 1505 and the session value column 1503 are derived from the household demographic characterization process 1400. The determinSC parameters such as address and telephone numbers can

be obtained from an outside source or can be loaded into the system by the subscriber or a network operator at the time of installation. Updating of deterministic values is prevented by indicating that these values should not be updated in the update column 1507.

FIG. 16 illustrates an entity-relationship diagram for the generation of the household interest profile 180 in a household interest profile generation process 1600. In a preferred embodiment, the household interest profile generation process comprises averaging the household session interest profile 1320 over multiple sessions and applying the household viewing habits 195 in combination with the heuristic rules 160 to form the household interest profile 180 which takes into account both the viewing preferences of the household as well as assumptions about households/subscribers with those viewing habits and program preferences.

FIG. 17 illustrates the household interest profile 180 which is composed of a programming types row 1709, a products types row 1707, and a household interests column 1701, an average value column 1703, and a session value column 1705.

The product types row 1707 gives an indication as to what type of advertisement the household would be interested in watching, thus indicating what types of products could potentially be advertised with a high probability of the advertisement being watched in its entirety. The programming types row 1709 suggests what kind of programming the household is likely to be interested in watching. The household interests column 1701 specifies the types of programming and products which are statistically characterized for that household.

As an example of the industrial applicability of the invention, a household will perform its normal viewing routine without being requested to answer specific questions regarding likes and dislikes. Children may watch television in the morning in the household, and may change channels during commercials, or not at all. The television may remain off during the working day, while the children are at school and day care, and be turned on again in the evening, at which time the parents may "surf" channels, mute the television during commercials, and ultimately watch one or two hours of broadcast programming. The present invention provides the ability to characterize the household, and may make the determination that there are children and adults in the household, with program and product interests indicated in the household interest profile 180 corresponding to a family of that composition. A household with two retired adults will have a completely different characterization which will be indicated in the household interest profile 180.

Although the present invention has been largely described in the context of a single computing platform receiving programming, the SCS 100 can be realized as part of a client-server architecture, as illustrated in FIG. 18. Referring to FIG. 18, residence 1800 contains a personal computer (PC) 1820 as well as the combination of a television 1810 and a set-top 1808, which can request and receive programming. The equipment in residence 1800, or similar equipment in a small or large business environment, forms the client side of the network as defined herein. Programming is delivered over an access network 1830, which may be a cable television network, telephone type network, or other access network. Information requests are made by the client side to a server 1840 which forms the server side of the network. Server 1840 has content locally which it provides to the subscriber, or requests content on behalf of the

subscriber from a third party content provider 1860, as illustrated in FIG. 18. Requests made on behalf of the client side by server 1840 are made across a wide area network 1850 which can be the Internet or other public or private network. Techniques for making requests on behalf of a client are frequently referred to a proxy techniques and are well known to those skilled in the art. The server side receives the requested programming which is displayed on PC 1820 or television 1810 according to which device made the request.

In a preferred embodiment the server 1840 maintains the subscriber selection data 110 which it is able to compile based on its operation as a proxy for the client side. Retrieval of source related information and the program target analysis process 1100, the program characterization process 800, the program target analysis process 1100, the session characterization process 1300, the household demographic characterization process 1400, and the household interest profile generation process 1600 can be performed by server 1840.

As an example of the industrial applicability of the invention, a household will perform its normal viewing routine without being requested to answer specific questions regarding likes and dislikes. Children may watch television in the morning in the household, and may change channels during commercials, or not at all. The television may remain off during the working day, while the children are at school and day care, and be turned on again in the evening, at which time the parents may "surf" channels, mute the television during commercials, and ultimately watch one or two hours of broadcast programming. The present invention provides the ability to characterize the household, and may make the determination that there are children and adults in the household, with program and product interests indicated in the household interest profile 180 corresponding to a family of that composition. A household with two retired adults will have a completely different characterization which will be indicated in household interest profile 180, which is stored at server 1840 and can be accessed by various groups wishing to provide programming and advertisements to the members of residence 1800.

Although this invention has been illustrated by reference to specific embodiments, it will be apparent to those skilled in the art that various changes and modifications may be made which clearly fall within the scope of the invention. The invention is intended to be protected broadly within the spirit and scope of the appended claims.

What is claimed is:

1. A method for generating a subscriber profile, the method comprising:
    monitoring subscriber interactions with a television;
    processing the subscriber interactions to generate viewing characteristics that identify viewing traits associated with the subscriber;
    retrieving heuristic rules associated with the viewing characteristics, wherein the heuristic rules predict traits about the subscriber not related to the viewing characteristics; and
    generating the subscriber profile by applying the heuristic rules to the viewing characteristics.

2. The method of claim 1, further comprising storing the subscriber profile.

3. The method of claim 1, wherein the heuristic rules are probabilistic in nature.

4. The method of claim 1, wherein the subscriber profile is probabilistic in nature.

5. The method of claim 1, wherein the heuristic rules predict demographic traits about the subscriber.

6. The method of claim 1, wherein the subscriber profile identifies demographic traits of the subscriber.

7. The method of claim 1, wherein the heuristic rules predict product interest traits about the subscriber.

8. The method of claim 1, wherein the subscriber profile identifies product interest traits of the subscriber.

9. The method of claim 1, wherein said monitoring includes monitoring at least some subset of channel changes, volume changes, record commands, and associated time.

10. The method of claim 1, wherein said processing includes aggregating the subscriber interactions to generate the viewing characteristics.

11. The method of claim 10, wherein said aggregating includes aggregating the subscriber interactions for a single viewing session to generate session viewing characteristics that identify viewing traits associated with the subscriber for that viewing session.

12. The method of claim 11, wherein
    said retrieving includes retrieving heuristic rules associated with the session viewing characteristics; and
    said generating includes generating a session subscriber profile by applying the heuristic rules to the session viewing characteristics.

13. The method of claim 10, wherein said aggregating includes aggregating the subscriber interactions for multiple viewing sessions to generate average viewing characteristics that identify viewing traits associated with the subscriber for the multiple viewing sessions.

14. The method of claim 13, wherein
    said retrieving includes retrieving heuristic rules associated with the average viewing characteristics; and
    said generating includes generating an average subscriber profile by applying the heuristic rules to the average viewing characteristics.

15. The method of claim 1, wherein the viewing characteristics do not identify raw subscriber interaction data.

16. The method of claim 1, wherein the viewing characteristics include at least some subset of
    viewing time per channel, category, and network;
    channel changes per time period;
    average volume per time period, channel, category, and network; and
    dwell time per channel, category, and network.

17. The method of claim 1, further comprising retrieving source related information corresponding to requested source material, wherein the source related information includes at least some subset of title, category, subcategory, description, rating, and content.

18. The method of claim 17, wherein said processing includes processing the subscriber interactions with respect to the source related information.

19. A method for generating a subscriber profile for a subscriber, the method comprising:
    monitoring subscriber viewing activities including requests for source material;
    retrieving source related information corresponding to requested source material, wherein the source related information includes descriptions of at least one aspect of the requested source material;
    creating a first representation of the subscriber based on the subscriber viewing activities and the source related information;
    retrieving a set of rules associated with at least a subset of information contained in the subscriber viewing activities and the source related information, wherein the set

of rules relates at least one aspect of the subscriber viewing activities or at least one aspect of the source related information to at least one non-viewing parameter; and

processing the subscriber viewing activities with respect to the source related information and the set of rules to generate the subscriber profile, wherein the subscriber profile defines at least a second representation of the subscriber.

20. The method of claim 19, wherein the set of rules define probabilities associated with the at least one non-viewing parameter and the subscriber profile is probabilistic in nature.

21. The method of claim 19, wherein at least one of the at least one non-viewing parameters is a demographic parameter and the subscriber profile defines at least a demographic representation of the subscriber.

22. The method of claim 19, wherein at least one of the at least one non-viewing parameters is a product interest parameter and the subscriber profile defines at least a product interest representation of the subscriber.

23. A method for generating a subscriber profile for a subscriber, the method comprising:

monitoring subscriber television viewing activities;

retrieving source related information corresponding to requested source material, wherein the source related information includes at least one description of the requested source material;

creating viewing characteristics based on the subscriber viewing activities and the source related information, wherein the viewing characteristics define a first representation of the subscriber;

retrieving a set of rules associated with at least some subset of the viewing characteristics, wherein the set of rules defines at least one aspect of the subscriber not defined in the viewing characteristics; and

generating the subscriber profile by applying the set of rules to the viewing characteristics, wherein the subscriber profile defines a second representation of the subscriber.

24. A system for generating a subscriber profile, the system comprising:

means for monitoring subscriber interactions with a television;

means for processing the subscriber interactions to generate viewing characteristics that identify viewing traits associated with the subscriber;

means for retrieving heuristic rules associated with the viewing characteristics, wherein the heuristic rules predict traits about the subscriber not related to the viewing characteristics; and

means for generating the subscriber profile by applying the heuristic rules to the viewing characteristics.

25. The system of claim 24, wherein said means for processing aggregates the subscriber interactions to generate the viewing characteristics.

26. The system of claim 24, further comprising means for retrieving source related information corresponding to requested source material, wherein the source related information includes descriptions of at least one aspect of the requested source material.

27. The system of claim 26, wherein said means for processing processes the subscriber interactions with respect to the source related information.

28. A system for generating a subscriber profile for a subscriber, the method comprising:

means for monitoring subscriber viewing activities including requests for source material;

means for retrieving source related information corresponding to requested source material, wherein the source related information describes at least one aspect of the requested source material;

means for creating viewing characteristics based on the subscriber viewing activities and the source related information, wherein the viewing characteristics define a first representation of the subscriber;

means for retrieving a set of rules associated with at least some subset of the viewing characteristics, wherein the set of rules defines at least one aspect of the subscriber not defined in the viewing characteristics; and

means for generating the subscriber profile by applying the set of rules to the viewing characteristics, wherein the subscriber profile defines a second representation of the subscriber.

29. The system of claim 28, wherein

said means for retrieving a set of rules retrieves a set of rules that define probabilities associated with the at least one aspect of the subscriber; and

said means for generating the subscriber profile generates the subscriber profile as a probabilistic determination of the at least one aspect of the subscriber not defined in the viewing characteristics.

30. The system claim 28, wherein

said means for retrieving a set of rules retrieves a set of rules that define demographic parameters of the subscriber; and

said means for generating the subscriber profile generates the subscriber profile as a demographic representation of the subscriber.

31. The system claim 28, wherein

said means for retrieving a set of rules retrieves a set of rules that define product interests of the subscriber; and

said means for generating the subscriber profile generates the subscriber profile identifying product interests of the subscriber.

32. A method for generating a subscriber profile, the method comprising:

monitoring subscriber interactions with a television;

processing the subscriber interactions in order to characterize interaction traits associated with the subscriber;

retrieving heuristic rules associated with the interaction traits, wherein the heuristic rules predict non-interaction traits about the subscriber, the non-interaction traits not being limited to child/adult distinctions; and

generating the subscriber profile by applying the heuristic rules to the interaction traits.

33. The method of claim 32, wherein said generating the subscriber profile includes generating the subscriber profile having at least one non-interaction trait that is not child/adult distinction related.

34. The method of claim 32, wherein the non-interaction traits are demographic traits.

35. The method of claim 32, wherein the heuristic rules associate the interaction traits to the non-interaction traits.

36. The method of claim 35, wherein interaction traits to non-interaction traits associations include at least some subset of

channel change speed to gender, income level or a combination thereof;

programs is to gender, work status, or a combination thereof; and

program categories, program sub-categories, program content or some combination thereof to gender, age, income level, family size, or some combination thereof.

37. The method of claim 32, wherein the heuristic rules assign probabilities for the subscriber having certain non-interaction traits associated therewith.

38. The method of claim 37, wherein the non-interaction traits having probabilities assigned thereto include at least some subset of age, income, family size and gender.

39. The method of claim 37, wherein each of the non-interaction traits are divided into categories and each category is assigned a probability.

40. The method of claim 32, wherein the subscriber profile assigns probabilities associated with the subscriber having certain non-interaction traits.

41. The method of claim 40, wherein the each of the non-interaction traits of the subscriber profile are divided into subcategories and each category is assigned a probability.

42. The method of claim 32, wherein the non-interaction traits are product interests.

43. The method of claim 42, wherein the product interests are not limited to products associated with children and products associated with adults.

44. The method of claim 32, wherein the subscriber profile may represent a single subscriber or a household of subscribers.

45. The method of claim 32, wherein the subscribers identity is not known.

46. A method for generating a subscriber profile, the method comprising:

processing subscriber interactions with a television in order to characterize interaction traits associated with the subscriber;

retrieving heuristic rules associated with the interaction traits, wherein the heuristic rules predict non-interaction traits about the subscriber; and

generating the subscriber profile by applying the heuristic rules to the interaction traits, wherein the subscriber profile is not limited to child adult distinctions, advertising associated with children or advertising associated with adults.

47. A method for generating a subscriber profile, the method comprising:

accessing interaction traits defining how a subscriber interacts with a television;

retrieving heuristic rules associated with the interaction traits, wherein the heuristic rules predict non-interaction traits about the subscriber, the non-interaction traits including at least some subset of gender, income level and family size; and

generating the subscriber profile by applying the heuristic rules to the interaction traits.

48. A method for generating a subscriber profile, the method comprising:

characterizing subscriber interactions with a television as interaction traits;

retrieving heuristic rules associated with the interaction traits, wherein the heuristic rules associate the interaction traits to non-interaction traits, the interaction traits to non-interaction traits associations including at least some subset of channel change speed to gender, channel change speed to income level, program genre to gender, program genre to income level and program genre to family size; and

generating the subscriber profile by applying the heuristic rules to the interaction traits.

* * * * *